Identification and characterization of gene defects underlying autosomal recessive intellectual disability in two Iranian families

Dissertation zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von
Lucia Püttmann
aus Oberhausen

2013

Die praktischen Arbeiten dieser Dissertation wurden vom 01.06.2008 bis 15.03.2013 unter der Leitung von Prof. Dr. Andreas W. Kuss und Dr. Luciana Musante am Max-Planck-Institut für molekulare Genetik, Abteilung Molekulare Humangenetik (Prof. Dr. H. H. Ropers) durchgeführt.

The work presented in this thesis has been conducted from 01.06.2008 until 15.03.2013 at the Max Planck Institute for Molecular Genetics, Department of Human Molecular Genetics (Prof. Dr. H. H. Ropers) under supervision of Prof. Dr. Andreas W. Kuss and Dr. Luciana Musante.

1. Gutachter: Herr Prof. Dr. Hans Hilger Ropers
2. Gutachter: Herr Prof. Dr. Stefan Sigrist

Disputation am 18.09.2013

"We have a hunger of the mind which asks for knowledge of all around us, and the more we gain, the more is our desire; the more we see, the more we are capable of seeing." Maria Mitchell, Astronomer (1818-1889)

Acknowledgements

First and foremost, I would like to thank Prof. Hans Hilger Ropers for giving me the opportunity to carry out my thesis at his department in a well-framed scientific environment and for his continuous support, guidance and encouragement.

I am grateful to Prof. Stephan Sigrist for being my second supervisor.

I whish to express sincere gratitude to Prof. Andreas W. Kuss and Dr. Luciana Musante for their guidance, support and encouragement throughout this study, and for helpful discussions and comments.

I would like to thank Prof. Hossein Najmabadi, Prof. Kimia Kahrizi and their colleagues at the Genetics Research Center, University of Social Welfare and Rehabilitation Sciences (Tehran, Iran) who have collected the clinical data that were used in this study. I also thank the patients and their families for their willingness to participate in this research.

I whish to thank all current and former lab members Agnes Zecha, Bettina Lipkowitz, Daniel Mehnert, Dr. Joanna Walczak Stulpa, Dr. Lars R. Jensen, Dr. Lia Abbasi Moheb, Marianne Schlicht, Dr. Masoud Garshasbi, Robert Weissmann, Sabine Otto, and Dr. Sahar Esmaeeli Nieh, who have created a pleasant working atmosphere and participated in helpful scientific discussions throughout. Special thanks go to Bettina Lipkowitz and Sabine Otto, who assisted me in some experimets.

Many thanks to Prof. Thomas Wienker, Dr. Andreas Tzschach and Dr. Hao Hu for their comments and profitable discussions.

I am also particularly grateful to Susanne Freier for technical assistance with cell culture and to Dr. Henning Stehr for *in silico* protein modeling and energy calculations.

Finally, I want to acknowledge my colleagues from the Department Human Molecular Genetics who have created a pleasant working atmosphere.

I am deeply grateful to my family, Victor Grönke and my friends for their continuous support.

Contents

1 Introduction

The human brain permanently processes and integrates information between specialized and functionally linked brain regions [van den Heuvel *et al.*, 2009]. Higher order cognitive processes are based on the capability of neurons to organize into networks and to remodel these networks during brain development as a response to learning and experience [Kramer and van Bokhoven, 2009]. Mutations in genes coding for components of these networks can lead to impairment of cognitive functioning with subtle to devastating effects on the mental health of the affected individuals. Elucidating the underlying genetic defects and studying their molecular impact on their respective gene products are therefore of great significance: understanding the molecular pathways and brain specific networks underlying normal intellectual performance provides the basis for expert genetic counseling of affected individuals and their families and guidelines for therapy.

1.1 Intellectual disability

Intellectual disabilitiy (ID), or mental retardation (MR), is defined by the World Health Organization, WHO, as a

> "significant reduced ability to understand new or complex information and to learn new skills (impaired intelligence). This results in a reduced ability to cope independently (impaired social functioning), and begins before adulthood, with a lasting effect on development [WHO, 2013]."

Individuals with an IQ of 70 or below are generally considered to be intellectually disabled [Ropers and Hamel, 2005]. According to the 10th International Classification of Diseases (ICD-10, Version 2007) six categories of intellectual disability exist, comprising mild, moderate, severe, profound, other and unspecified intellectual disability (Table 1.1) [WHO, 2010].

To support or supplement the diagnosis of ID, skills in adaptive behaviour can be measured by standardized tests (e. g. Wechsler Intelligence Scale for Children or for Adults [Wechsler, 2004, 2008]. According to the American Association on Intellectual and Developmental Disabilities (aaidd) these tests assess the following three skill types:

1. conceptual skills, comprising self-direction, language and literacy as well as money, time, and number concepts;

2. social skills, including interpersonal skills, social responsibility, self-esteem, social problem solving, and following rules;

Table 1.1: Degrees of intellectual disability based on ICD-10, version 2010

Category	IQ	Mental age	Level of support and self-care
Mild	50 – 69	9 – < 12 years	likely to result in some learning difficulties in school; many adults will be able to work, maintain good social relationships and contribute to society
Moderate	35 – 49	6 – < 9 years	likely to result in marked developmental delays in childhood but most can learn to develop some degree of independence in self-care and acquire adequate communication and academic skills; adults will need varying degrees of support to live and work in the community
Severe	20 – 34	3 – < 6 years	likely to result in continuous need of support
Profound	< 20	< 3 years	results in severe limitation in self-care, continence, communication and mobility
Other	–	–	this category is only used if determination of the ID degree is difficult or impossible due to associated sensory or physical impairments, as in blind, deaf-mute, and severely behaviourally disturbed or physically disabled individuals
Unspecified	–	–	evidence of ID in the affected individual, but insufficient information is available to assign the patient to one of the above categories

3. practical skills, enompassing e. g. activities of daily living, occupational skills, and healthcare.

Intellectual disability can also be divided into syndromic and non-syndromic forms. In syndromic forms affected individuals present with additional clinical features resulting in a distinct phenotype such as e. g. in Down Syndrome. Non-syndromic ID (NSID) is defined as ID being the exclusive clinical finding [Kaufman *et al.*, 2010]. Diagnosis of NSID can be challenging and sometimes misleading, as syndromes can be subtle and only be assessed clinically after previous genetic diagnosis [Ropers, 2006].

The prevalence of ID in Western countries is 1.5 to 2 % and severe ID affects 0.3 to 0.5 % of the population [Ropers, 2008, 2010a]. ID can be of environmental and/or genetic origin. Environmental factors comprise exposure to teratogens, viruses or radiation as well as lack of oxygen in the brain caused by head injury [Kaufman *et al.*, 2010]. However, the vast majority of ID, especially severe ID, is thought to be due to genetic factors [McLaren and Bryson, 1987] such as large chromosomal abberations and rearrangements, copy number variations (CNVs) or single gene defects.

Numerical or structural chromosomal abberations are found in one out of seven individuals with severe ID [Leonard and Wen, 2002]. Chromosomal abberations that are detectable under the microscope generally give rise to a syndromic phenotype, as they are associated with large genetic imbalances affecting a great number of genes [Pfundt and Veltman, 2012]. The most common chromosomal abberation with a total prevalence of 22 per 10 000 births in Europe is Down Syndrome (trisomy 21), which is further the most commonest congenital cause of intellectual disability [Leonard and Wen, 2002]. Trisomies are usually lethal at embryonic or fetal stages. Less common trisomies that survive to term are Pätau syndrome (trisomy 13) and Edwards syndrome (trisomy 18) with total prevalence of 5 and 2 per 10 000 births in Europe, respectively [Loane *et al.*, 2012]. The vast majority of chromosomal abberations are unbalanced. It has been estimated that truly balanced chromosomal rearrangements account for disease in less than 1% of affected individuals with ID [Hochstenbach *et al.*, 2009]. A copy number variant (CNV) is defined as a DNA segment that is 1 kb or larger and present at variable copy numbers in comparison with a reference genome [Redon *et al.*, 2006; Feuk *et al.*, 2006].

Copy number variations (CNVs) encompass deletions, insertions, duplications, and complex multi-site variants [Fredman *et al.*, 2004]. CNVs that are associated with ID often contain several genes that are expressed specifically in the brain or are enriched in neurodegenerative pathways [Webber *et al.*, 2009]. For many microdeletion syndromes it has yet to be identified which gene in the common deleted region is responsible for which phenotypic feature present in the syndrome [Raymond and Tarpey, 2006].

Single gene defects give rise to a large variety of X-linked or autosomal ID forms as introduced in the following sections. Disease-causing alterations underlying monogenic disorders encompass a wide range of mutation categories. Small insertions or deletions and splice site mutations can lead to frameshifts that alter the translational reading frame of mRNA. Mutations that result in aberrant splicing can be caused by mutations that change the conserved intron donor and accpetor sequences, but also by mutations in other cis-acting regulatory sites or by mutations that activate cryptic splice cites. Nonsense mutations generate premature stop codons and can give rise to truncated proteins, but most nonsense mutations result in null alleles as they trigger nonsense-mediated mRNA decay (NMD). Single gene defects that give rise to proteins comprise missense mutations and in-frame insertions and deletions [Strachan and Read, 2003].

1.1.1 X-linked intellectual disability

The X chromosome is 155 Mbp long and contains approximately 5 % of the haploid human genome [Ross *et al.*, 2005]. In males up to 12 % of ID is caused by mutations in X chromosomal protein coding genes [Ropers, 2010a,b, 2008; Ropers and Hamel, 2005]. As males only carry the maternally derived X chromosome, any mutation in genes without counterpart on the Y chromosome will have a direct impact on the phenotype. Females inherit one X chromosome from each parent, but only one is active, whereas the other one is randomly inactivated and the outcome of X-linked intellectual disability (XLID) is dependent on the percentage of cells in which

the mutated allele is expressed [Chelly and Mandel, 2001].

More than 500 genes located on the X chromosome are expressed in the brain. They encode transcription factors, channels, receptors, DNA- and RNA-binding proteins, scaffolders, enzymes, and components involved in signal transduction [Laumonnier *et al.*, 2007]. X chromosomal genes seem to influence general intelligence as well as social cognition and emotional regulation. XLID candidate genes function in cell migration, development and regulation of neuronal networks, cell to cell communication and brain development [Ross *et al.*, 2005; Lisik, 2010].

Fragile X syndrome

The fragile X syndrome (FXS, MIM #300624) accounts for up to 25 % of all XLID cases and thus is the most abundant form of XLID. It has a prevalence of 1:4000 – 1:6000 in males and 1:8000 – 1:10 000 in females. FXS is characterized by ID, behavioural problems, distinctive facial features and features characteristic for autism [Hagerman *et al.*, 2012]. The variable phenotype observed in individuals with FSX can be due to germline mosaicism and, in the female patients, to X-inactivation [Bhakar *et al.*, 2012].

In the vast majority of cases FSX is caused by a CCG trinucleotide extension in the 5'UTR of *FMR1*, that leads to hypermethylation and thus transcriptional silencing, resulting in the absence of the gene product, FMRP (fragile X mental retardation protein). Known point mutations affecting *FMR1* and resulting in FXS comprise two missense mutations (p.I304N [De Boulle *et al.*, 1993] and p.R138Q [Collins *et al.*, 2010]), a 1-bp deletion in exon five leading to a frameshift and premature stop-codon, and a 2-bp change affecting the intron/exon boundary of the second exon [Lugenbeel *et al.*, 1995]. In addition, a microduplication encompassing *FMR1* and *ASFMR1* was found to result in developmental delay, epilepsy and hyperactivity [Vengoechea *et al.*, 2012].

Affected individuals as well as *Fmr1* knockout mice show an increase in dendritic spine density and abnormal spine morphology [Comery *et al.*, 1997; Hinton *et al.*, 1991; Irwin *et al.*, 2000]. Moreover, proteins involved in synaptic functioning and neuronal activity are decreased in the prefrontal cortex of *Fmr1* knockout mice [Krueger *et al.*, 2011].

Several treatment strategies for FSX have been developed. Lithium reverses synaptic plasticity deficits and increases rates of cerebral protein synthesis in *Fmr1* knockout mice [Choi *et al.*, 2011; Liu *et al.*, 2012]. In affected human individuals, lithium improved behaviour as well as adaptive and cognitive skills in a pilot trial [Berry-Kravis *et al.*, 2008]. Several FXS symptoms concerning the nervous system are caused by unbalanced activation of group one mGluRs (metabotropic glutamate receptors) [Dölen *et al.*, 2007]. FMRP and mGluR5 act as antagonists. Fenobam, a selective mGluR5 antagonist [Porter *et al.*, 2005], improves behaviour and cognitive traits [Berry-Kravis *et al.*, 2009]. Several selective mGluR5 negative allosteric modulators are currently being tested in clinical trials for FXS (www.clinicaltrials.gov).

Other forms of XLID

To date, at least 230 different XLID disorders, including 160 syndromic forms have been reported. So far, at least 72 genes implicated in syndromic XLID have been cloned, of which 19 genes give rise to non-syndromic XLID and 17 genes have been implicated in both kinds of XLID [Lubs *et al.*, 2012; Lisik, 2010]. These results underline that XLID is very heterogeneous [Ropers, 2006].

Major functions of genes associated with XLID are involved in transciptional regulation (approx. 22 %), signal transduction (approx. 19 %), metabolism (approx. 15 %) or have membrane-associated functions (approx. 15 %). The remaining roles (each with about 3–5 %) of XLID genes are in DNA- and RNA processing, DNA metabolism, protein synthesis, ubiquitination, cytoskeleton, cell cycle and cell adhesion [Ropers, 2008].

Following FSX as the chief cause for XLID, the second most common cause of XLID (5 %) are defects in *ARX* (aristaless-related homeobox gene; for review see e. g. [Gécz *et al.*, 2006; Shoubridge *et al.*, 2010]). Mutations in *ARX* cause a large spectrum of brain disorders (Mental retardation, X-linked 36/43/54 MIM #300419; Epileptic encephalopathy, early infantile, 1 MIM #308350; Lissencephaly, X-linked 2 MIM #300215; Hydranencephaly with abnormal genitalia MIM #300215; Partington syndrome MIM #309510 and Proud syndrome MIM #300004). As a member of the Aristaless-related gene family encoding transcription factors, ARX is required for crucial steps during development of the central nervous system (CNS) [Kaufman *et al.*, 2010]. ARX target genes control cortical interneuron migration and differentiation [Friocourt and Parnavelas, 2011]. *ARX* mutation analyses, including mutations present in ARX disorders, have provided evidence that mutations located in the homeodomain result in a loss of DNA binding activity and in a loss of transcriptional repression activity [Cho *et al.*, 2012; Shoubridge *et al.*, 2012]. Further genes prevalent in XLID include *JARID1C*, *SLC6A8*, and *CUL4B* [Ropers, 2008]. For in-depth review see e. g. [Lubs *et al.*, 2012; Gécz *et al.*, 2009].

1.1.2 Autosomal intellectual disability

Autosomal dominant intellectual disability

Autosomal dominant (AD) forms of severe ID are without much doubt almost exclusively the result of *de novo* mutations, as affected individuals will usually stay without offspring. The prevelance of dominant ID is currently unknown, but ADID is thought to be quite common [Ropers, 2010b], since *de novo* point mutations, for example, occur at a rate of about 2.0×10^{-8} per base pair per generation [O'Roak *et al.*, 2012; Sanders *et al.*, 2012; Iossifov *et al.*, 2012]. Fifteen genes have been associated with non-syndromic ADID to date (Table 1.2).

The most common cause of non-syndromic ADID are truncating mutations in *SYNGAP1*, which have also been associated with autism [Hamdan *et al.*, 2009, 2011a]. *SYNGAP1* encodes a neuronal Ras GTPase activating protein that lies downstream of NMDA receptors and is an intermediate required for neural circuit function and behaviour [Guo *et al.*, 2009]. Interaction of SYNGAP1 with PSD-95, a scaffolding protein of the postsynaptic density, is required for deter-

Table 1.2: ADID genes and their molecular function

Gene	Protein function	Reference
ARID1B	component of the SWI/SNF chromatin remodeling complex	[Hoyer *et al.*, 2012; Santen *et al.*, 2012; Tsurusaki *et al.*, 2012]
CACNG2	type I transmembrane AMPA receptor regulatory protein (TARP)	[Hamdan *et al.*, 2011b]
CDH15	calcium-dependent intercellular adhesion glycoprotein	[Bhalla *et al.*, 2008]
CIC	transcriptional repressor	[Vissers *et al.*, 2010]
DEAF1	transcription factor	[Vissers *et al.*, 2010]
DOCK8	guanine nucleotide exchange factor	[Griggs *et al.*, 2008]
DYNC1H1	dynein	[Vissers *et al.*, 2010]
EPB41L1	multifunctional protein that mediates interactions between the erythrocyte cytoskeleton and the overlying plasma membrane	[Hamdan *et al.*, 2011b]
GRIN1	subunit of N-methyl-D-aspartate receptors (glutamate receptor)	[Hamdan *et al.*, 2011b]
MDB5	methyl-CpG binding domain protein 5	[Wagenstaller *et al.*, 2007]
KIF1A	kinesin	[Hamdan *et al.*, 2011b]
KIRREL3	nephrin-like protein	[Bhalla *et al.*, 2008]
SHANK2	scaffold protein in the postsynaptic density	[Berkel *et al.*, 2010]
SYNGAP1	synaptic Ras GTPase activating protein 1	[Hamdan *et al.*, 2009]
YY1	transcription factor	[Vissers *et al.*, 2010]

mining the timing of spine formation and the size of mature spines [Vazquez *et al.*, 2004]. SYNGAP1 is also involved in the regulation of spine morphology and in transient NMDA-receptor-dependent regulation of the spine cytoskeleton [Carlisle *et al.*, 2008]. Homozygous *SynGAP* $^{(-/-)}$ mice die shortly after birth [Komiyama *et al.*, 2002; Kim *et al.*, 2003; Vazquez *et al.*, 2004]. Adult heterozygous *SynGAP* $^{(+/-)}$ mice exhibit a reduced amplitude of hippocampal long-term potentiation (LTP) and impaired learning in the Morris Water Maze [Komiyama *et al.*, 2002]. It has been found that a reduction of SynGAP concentration starts during the first week of life, leading to enhanced neuronal apoptosis [Knuesel *et al.*, 2005].

Autosomal recessive intellectual disability

In consanguineous populations recessively inherited congenital disorders including autosomal recessive forms of ID (ARID) are up to ten times more frequent than in outbred populations [Modell and Darr, 2002], reflecting the high proportion of homozygous DNA in offspring of consanguinous parents. Parental consanguinity and large families are common in Nothern Africa and the Middle East. In Iran, for example, 40 % of children are born to consanguineous parents [Najmabadi *et al.*, 2011]. This leads to an increased inheritance of recessive traits from

both parents (autozygosity, i. e. homozygosity for alleles identical by descent) [Lander and Botstein, 1987; Kaufman *et al.*, 2010]. The strategy of choice for investigating molecular causes of ARID is a combination of linkage analysis and homozygosity mapping in large consanguineous families followed by mutation screening in thus identified candidate genes [Bull *et al.*, 1998; Lander and Botstein, 1987]. So far, all of the genes implicated in ARID have been identified using this approach. All but the seven following genes have been reported in a single family:

ELP2 ELP2 is part of the RNA polymerase II elongator complex. This complex is a histone acetyltransferase component of RNA polymerase II. ELP2 might be involved in chromatin remodeling and has a role in acetylation of histones H3 and H4. Defects in *ELP2* are present in two families with NSID [Najmabadi *et al.*, 2011].

MAN1B1 ERManI (endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase), encoded by *MAN1B1*, is a member of the glycosyl hydrolase family 47 (GH47). These enzymes are involved in N-glycan processing and also in endoplasmic reticulum quality control, preparing misfolded glycoproteins for degradation [Karaveg *et al.*, 2005]. To date, five families have been identified with defects in *MAN1B1* segregating with mild to moderate syndromic ID (SID) or NSID [Rafiq *et al.*, 2011; Najmabadi *et al.*, 2011].

NSUN2 NSUN2 (NOP2/Sun domain family, member 2) exhibits methyltransferase activity against hemimethylated DNA as well as rRNA and tRNA *in vitro* [Frye and Watt, 2006]. Furthermore, the protein interacts with the mitotic spindle and is involved in cell division. The latter functions are independent of the methyltransferase activity of NSUN2 [Hussain *et al.*, 2009]. Deletion of the *NSUN2* ortholog in *D. melanogaster* leads to severe impairment of short-term-memory [Abbasi-Moheb *et al.*, 2012]. Five families present with mutations in *NSUN2* and moderate to severe SID. *NSUN2* has recently been implicated in a Dubowitz-like syndrome encompassing mild microcephaly, ID, growth retardation, facial dismorphologies and eczema [Martinez *et al.*, 2012].

ST3GAL3 *ST3GAL3* encodes the β-galactoside-α2,3-sialyltransferase-III (ST3Gal-III). This Golgi enzyme catalyses the formation of the epitope sialyl Lewis^a (which is involved in cell-to-cell recognition) on glycoproteins. In affected individuals ST3Gal-III activity is severly impaired and the protein displays a cellular destribution in the endoplasmic reticulum. Defects in *ST3GAL3* are present in two families and the affected individuals suffer from mild to severe NSID [Hu *et al.*, 2011].

TRAPPC9 The most common cause for ARID today are mutations in *TRAPPC9*. In six families the patients suffer from moderate to severe SID [Mir *et al.*, 2009; Philippe *et al.*, 2009; Mochida *et al.*, 2009; Abou Jamra *et al.*, 2011; Marangi *et al.*, 2012].

The gene *TRAPPC9* encodes subunit nine of the transport protein particle complex (TRAPP), a conserved protein complex involved in endocytic and secretory pathways [Sacher *et al.*, 2008]. TRAPPC9 functions in nerve growth factor-induced differentiation and enhances TNFα-induced NF-κB activation *in vitro* [Hu *et al.*, 2005]. Loss of *TRAPPC9* in affected individuals might disrupt neuronal differentiation [Mir *et al.*, 2009].

TUSC3 TUSC3 has a vital function in the vertebrate plasma membrane magnesium ion transport system [Zhou and Clapham, 2009]. As in rats increased magnesium ion levels in brain enhance learning and memory, imbalanced magnesium ion transport caused by loss of *TUSC3* is probably the cause of ID in affected individuals [Garshasbi *et al.*, 2011]. Defects in *TUSC3* have been identified in four kindreds with ID ranging from moderate to severe [Garshasbi *et al.*, 2008; Molinari *et al.*, 2008; Garshasbi *et al.*, 2011; Khan *et al.*, 2011].

ZNF526 ZNF526 is a krüppel-type zinc finger protein involved in transcription regulation [SwissProt Accession Number Q8TF50]. Mutations in *ZNF526* lead to severe NSID in two kindreds [Najmabadi *et al.*, 2011].

Mutations causing ARID affect a large variety of functionally different genes, including e. g. housekeeping genes coding for histones, histone demethylases, proteins involved in transcriptional regulation, splicing, protein degradation, fatty acid synthesis and turnover, cell cycle control, and cell migration. ARID genes with brain specific functions are involved in glia cell differentiation, regulation of neurotransmission, calcium channel functioning, exocytosis or neurotransmitter release as well as components of the Ras and Rho signaling pathways, that are critically involved in synapse function and formation as well as during neurodevelopment. It is obvious that ARID is extremely heterogenous, even much more than XLID, and common forms apparently do not exist. Furthermore, the majority of genes identified in ARID are ubiquitously expressed and have vital functions involving basic cellular processes rather than being confined to neuron or synapse function [Najmabadi *et al.*, 2011].

2 Aim of this study

Early onset ID is the most frequent handicap in children and the vast majority of ID caused by single gene defects is due to mutations on the autosomes. This thesis aims at the the identification and functional characterization of specific genetic defects in two Iranian families with ARID, where systematic clinical studies and autozygosity mapping was performed before during a collaborative project between the Max Planck Institute for Molecular Genetics, Berlin, Germany and the Genetics Research Center (GRC) at the University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

3 Patients, Materials and Methods

3.1 Patients

This study was carried out in accordance with the ethical standards of the appropriate national and institutional committees. Affected individuals and their families were recruited by the Genetics Research Center (GRC) at the University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. In all cases, informed consent was given by healthy family members and the patient's parents. Photographs were taken to determine physical findings and the mental status of the affected family members was examined using a modified version of the Wechsler Intelligence test. The karyoptype of each patient was examined and fragile X-syndrome (FXS) was excluded.

3.1.1 Family M289

Family M289 is an Iranian family with arabic ethnicity and four individuals presenting with ARID (Figure 3.1). The affected family members presented with moderate ID and borderline microcephaly with the occipitofrontal circumference (OCF) below the third centile (Table 3.1). The patients and the parents' height are below the third centile as well. DNA was extracted from peripheral blood of the patients (V:1, V:2, V:4, V:6), three healthy siblings (V:3, V:5, V:7) and their parents (IV:1 and IV:2) using standard procedures. The karyoptype of all patients was normal.

Table 3.1: Clinical features of family M289

Family member	Sex	Age at examination	IQ	Height	OFC	Additional features
V:1	m	29	45	158	51	pes platus,thin body
V:2	m	12	40-50	151	48	pes platus,thin body
V:4	f	16	45	151	48	NS
V:6	f	14	40	151	48	NS
IV:1	f	–	–	151	53	–
IV:2	m	–	–	165	55	–

OCF occipitofrontal circumference; NS not seen

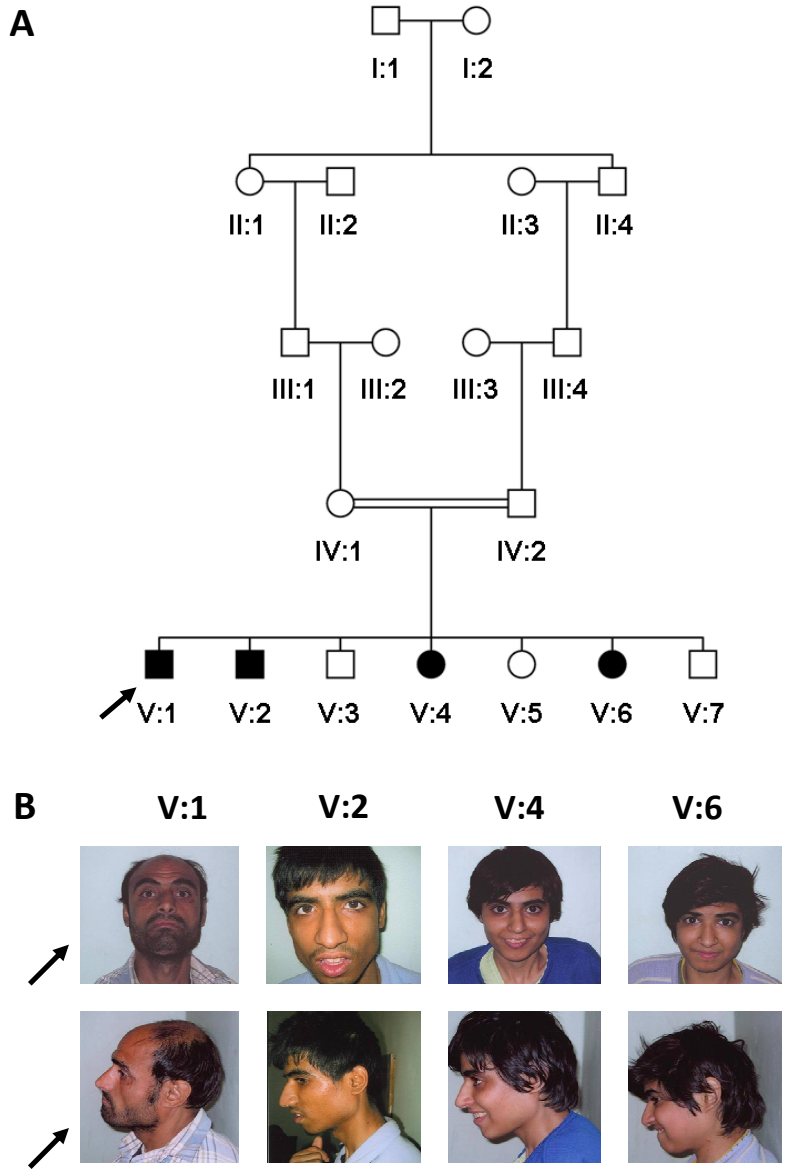

FIGURE 3.1: **A**) Pedigree of family M289. Full symbols denote affected individuals with ID. **B**) Facial images of affected individuals. V:1 (arrow) is the index patient.

3.1.2 Family 8600485

Family 8600485 is an Iranian family with Fars ethnicity and four individuals presenting with severe ID and a history of generalized tonic-clonic seizures without hyperactivity or attention deficit (Figure 3.2). All affected individuals were born full-term after uneventful pregnancies with birth weights between 3000 and 3500 grams. During the neonatal period all affected individuals suffered from hypotonia. All patients had developmental delay (Table 3.2). They started to speak at four to six years of age and walking had not started until about age 3.5. Patient V:9 (Figure 3.2) has a history of unilateral glaucoma, which has started at the age of 11 years. Apart from epilepsy, the patients suffered from no other neurologic disorders and they had no neuropsychiatric problems. Seizures were successfully treated with carbamazepine. Brain magnetic resonance imaging (MRI) scans or metabolic tests were not performed in this family.

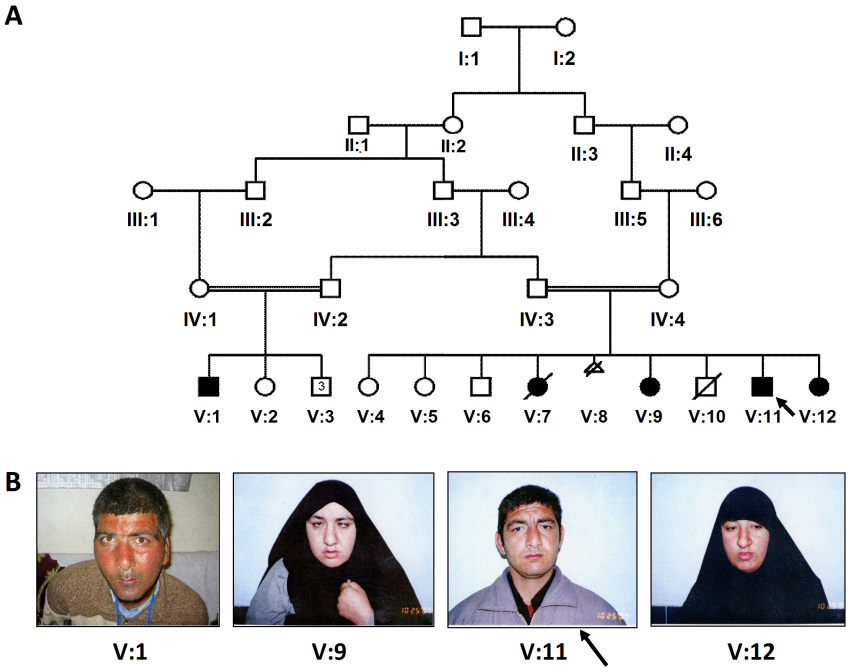

FIGURE 3.2: **A**) Pedigree of family 8600485. Full symbols denote affected individuals with ID. **B**) Facial images of affected individuals. V:11 (arrow) is the index patient.

Table 3.2: Clinical features of the affected members of family 8600485

Patient	Sex	Age at examination	IQ	Height	OFC
V:1	m	27	–*	181	55.0
V:9	f	27	30	163	51.1
V:11	m	23	35	181	55.5
V:12	f	21	35	157	52.0

OCF occipitofrontal circumference; * diagnosed as severely intellectually disabled without IQ testing

3.2 Materials

3.2.1 Consumables (disposable materials)

Product	Supplier
ABgene® PCR Plates	Thermo Scientific
Adhesive PCR Film	Thermo Scientific
BD Plastipak®, 50 ml	Luer
BD Discardit™ II, 10 ml	Luer
Cellstar® tubes, 50 ml	Greiner Bio-One
Cell culture flasks	TTP
Combitips® PLUS 0.1 and 0.2 ml	Eppendorf
Cover slips	Menzel-Gläser
Cryo.S™ with srew cap	Greiner Bio-One
Disposable reaction tube 14 ml	Greiner Bio-One
Disposable reaction tube 30 ml	Sarstedt
EIA/RIA 96 Well, no lid, flat bottom, high binding, costar	Corning
Falcon tube	Greiner Bio-One
FIA-Plate, black, 96 Well, flat bottom, non binding	Greiner Bio-One
Flexible Plate, 96 Well, U Bottom without Lid	BD Falcon™
Hypercassette™	Amersham
MicroAmp™ 8-Cap Strip	Applied Biosystems
Microscopic slides	Roth
MµltiFlex Round Tips	Sorenson BioScience
MµltiFlex Ultra PCR tube	Sorenson BioScience
Non pyrogenic serological pipette (2, 5, 10 & 25 ml)	Corning
Parafilm	Pechiney Plastic Packaging
Pasteur pipette	Roth
PCR plate (96 well)	ABgene
Premium Tips	Biozym
PVDF Western Blotting Membranes	Roche
Reaction tube (1.5 & 2.0 ml)	Eppendorf
Röhre 30 ml	Sarstedt
Rotilabo®-Spritzenfilter, PVDF, steril 45 µm	Roth
Safe-Lock Tubes, 0.5, 1.5 and 2.0 ml	Eppendorf
Surgical Disposable Scalpel	Aesculap
Space Saver® 10 µl	Rainin
Super RX FUJI Medical X Rax Film	Fujifilm
Thermowell® Gold PCR Plates	Corning
Thermowell® 96 Well PCR Plates	Corning
Ultra Tube, 0.65 ml	Roth
UVette® 220-1600 nm	Eppendorf
Whatman paper	Sigma

3.2.2 General reagents

Chemical product	Supplier
Acetic acid	Merck
ATP disodium salt hydrate	Sigma
Albumin, from bovine serum	Sigma
Ammonium persulfate	Sigma
Ampicillin	Sigma
Aqua ad iniectabilia	Baxter
β-mercaptoethanol	Whatman
Bacto™ Tryptone	BD Biosciences
Bacto™ Yeast Extract	BD Biosciences
Betaine	Sigma
BigDye®	Applied Biosciences
BigDye® Terminator	Applied Biosciences
BIOMOL GREEN Reagent™	Enzo life sciences
Bio-X-Act™ Long Mix	Bioline
Bromophenol Blue	Sigma
Chloramphenicol	Sigma
Chloroform	Merck
Complete, Mini Protease Inhibitor Cocktail Tablets	Roche
DAPI	Serva
DEPC	Aldrich
Digitonin	Sigma
DTT	Invitrogen
DMEM	Lonza
dNTPs	Roth
EDTA	Merck
Ethanol	Merck
Ethidiumbromide solution	Sigma
Expand Long Range dNTP Pack	Roche
FailSafe™ PCR System	Epicentre Biotechnologies
FBS	Sigma
First strand buffer 5 ×	Merk
Ficoll™	Amersham Biosciences
Fluoromount G	Southern Biotechnologies
Formaldehyde – 37 % (v/v)	FlukaBiochemika
Formamide	FlukaBiochemika
Glycerol – 85 % and 100 %	Merck
Glycin	Merck
HEPES	USB Corporation
Imperial™Protein Stain	Thermo Scientific
IPTG	Roth
Isopropanol	Merck

Chemical product	Supplier
Hydrochloric acid	Merck
Kanamycin	Invitrogen
LipoFECTAMINE™2000 Reagent	Invitrogen
L-glutamine	Lonza
L-serine	Sigma
Magnesium chloride	Merck
Methanol	Merck
Milk powder	Uelzena eG
NAD^+	Sigma
NADH	Sigma
OPTI-MEM®	Life Technologies
P/S	Lonza
PFA	Merck
Phenol	Sigma
Phenol : chloroform : isoamyl alcohol (25 : 24 : 1)	Sigma
Potassium dihydrogen phosphate	Merck
Potassium chloride	Sigma
Random Primers	Promega
Rotiphorese® Gel 30	Roth
SDS	Bio-Rad
Sodium acetate	Sigma-Aldrich
Sodium chloride	Roth
Sodium hydroxide	Merck or Sigma
Sodium hydrogen phosphate	Merck
Succinic semialdehyde	Santa Cruz Biotechnology
TEMED	Gibco BRL
Triton® X-100	Sigma-Aldrich
TRIzol reagent	Gibco BRL
Trypsin/EDTA solution	Lonza
Tween® 20	Sigma-Aldrich
UltraPure™ Agarose	Invitrogen
Urea	Biorad
Xylencyanol	Roth
X-Gal	Roth

3.2.3 Kits and Markers

Product	Supplier
1 kbp DNA ladder	Roth
GeneChip® Human Mapping 50K Array and Assay Kit	Affymetrix
GST Bulk Kit	GE-Healthcare
Human610-Quad BeadChip Kit	Illumina
HyperLadder™ I	Bioline
HyperLadder™ IV	Bioline
HyperLadder™ V	Bioline
MinElute PCR Purification Kit	Qiagen
QIAfilter™ Plasmid Maxi Kit	Qiagen
QIAprep® Spin Miniprep Kit	Qiagen
QIAquick® Gel Extraction Kit	Qiagen
QIAshredder	Qiagen
PageRuler™ Prestained Protein Ladder	Fermentas
PD-10 Columns	GE-Healthcare
Phase Lock Gel Heavy, 2 ml	Eppendorf
pUC Mix Marker, 8	Fermentas
QIAquick Gel Extraction Kit	Qiagen
QuickChange™ II XL Site-Directed Mutagenesis Kit	Stratagene

3.2.4 Equipment

Miscellaneous	
Name	**Supplier**
Concentrator 5301	Eppendorf
LSM 700 (confocal microscope)	Zeiss
Microwave privileg 9023G	Privileg
NanoDrop® ND-1000 Spectrophotometer	Peqlab
pH meter 766 Calimatic	Knick
POLARstar Omega (microplate reader)	BMG Labtech
Promax 2020 (shaker)	Heidolph
Research pro 100 electropipette	Eppendorf
Sonopuls HD 2070 (ultrasonic cell disruptor)	Bandelin
Ultrospec II 4050 (photometer)	LKB Biochrom
Water bath	Köttermann
Electrophoresis tanks	
Name	**Supplier**
Electrophoresis System	GibcoBRL
Horizon® 11-14 and 20-25	Heraeus
Horizontal Gel Incubator B 5050 E	Heraeus
Maxi-Gel-Kammer HU25, No. 2562	Roth

Centrifuges	
Name	**Supplier**
Centrifuge 5417R	Eppendorf
Centrifuge 5810R	Eppendorf
Rotanta 46K	Hettich
Sorvall RC-5 and Sorvall RC-5B Refrigerated Superspeed Centrifuge	Dupont Instruments-Sorvall
Incubators	
Name	**Supplier**
Incubator shaker C24	New Brunswick Scientific
Incubator shaker G25	New Brunswick Scientific
Forma Scientific CO_2 water jacketed incubator 3121	Thermo Scientific
PCR cyclers	
Name	**Supplier**
DNA Engine Tetrad Peltier Thermal Cycler	Bio-Rad
GeneAmp® PCR System 9700	Applied Biosystems
Peltier Thermal Cycler PTC-225	MJ Research
Name	**Supplier**
MR 3002	Heidolph
MR 1000	Heidolph
Thermomixer 5436	Eppendorf
Thermomixer comfort 1.5 ml	Eppendorf
Western Blotting Equipment	
Name	**Supplier**
Mini-PROTEAN® 3 Cell SDS-PAGE system	Bio-Rad
PowerPac HC Power Supply	Bio-Rad
PowerPac Universal Power Supply	Bio-Rad
Trans-Blot® SD Semi-Dry Transfer Cell	Bio-Rad

3.2.5 Buffers, solutions and media

Solution	Composition
Ampicillin	50 mg/ml ampicillin was dissolved in ddH_2O, sterilized by filtration and stored at −20 °C
APS 10	10 % w/v APS was dissolved in ddH_2O, aliquoted and stored at −20 °C
Blocking buffers	a) 3 % BSA in 1× PBST b) 5 % milk powder in 1× PBST
Miniprep buffer 1	50 mM glucose

Solution	Composition
	10 mM EDTA 25 mM Tris-HCl, pH 8.0
Miniprep buffer 2	0.2 M NaOH 1 % SDS
Miniprep buffer 3	3 M sodium acetate, pH 4.8
Cell Fractionation Lysis Buffer	20 mM HEPES, pH 7.4 50 mM NaCl 5 mM $MgCl_2$ 16 µl 25 × complete protease inhibitor solution 80 mg/ml digitonin
Chloramphenicol	34 mg/ml chloramphenicol was dissolved in ethanol and stored at −20 °C
DNA re-suspension buffer	0.4 M Tris-HCl, pH 8.0 0.06 M NaEDTA 0.15 M NaCl
DNA-loading buffer	20 % Ficoll 0.1 % Bromophenol blue 0.1 % Xylencyanol in ddH_2O
Ethidium bromide	10 mg/ml EtBr was dissolved in ddH_2O
Elution buffer	50 mM Tris-HCl, pH 8.0 10 mM Glutathione
Kanamycin	10 mg/ml kanamycin was dissolved in ddH_2O, sterilized by filtration and stored at −20 °C
Mager Mix	50 µl dNTP Mix 5 ml 10 × PCR buffer 5 ml 25 mM $MgCl_2$ 5 ml 5 M betaine 34.3 ml ddH_2O 500 µl MPI-Taq 10 U/µl 14 µl MPI-Pfu-Taq 10 U/µl

Solution	Composition	
Magic Mix	48 %	urea
	15 mM	Tris-HCl, pH 7.5
	8.7 %	glycerol
	1 %	SDS
	0.004 %	Bromophenol Blue
	143 mM	β-mercaptoethanol
PBS 1× buffer	137 mM	NaCl
	2.7 mM	KCl
	10.1 mM	Na_2HPO_4
	1.8 mM	KH_2PO_4
PBST 1× buffer		1× PBS
	1:1000	Tween 20
Pyrophosphate release reaction buffer	100 mM	Tris-HCl, pH 7.6
	10 mM	$MgCl_2$
	40 mM	KCl
	1 mM	DTT
	1 mM	ATP
	0.2 mM	L-serine
SDS-lysis buffer	0.1 M	Tris-HCl, pH 6.8
	0.1 M	NaCl
	1 M	β-mercaptoethanol
	5 %	SDS
	15 %	glycerol
SDS-PAGE electrode buffer 1×	72 g	glycine
	25 ml	10 % SDS
	15 g	Tris
	add 5 l	ddH_2O
Sepharose column binding buffer	10 mM PBS, pH 7.4	
SSADH-buffer	100 mM	Tris-HCl, pH 8.6
	50 mM	KCl
	0.1 mM	EDTA
	20 mM	*β*-mercaptoethanol
	0.1 %	Triton® X-100

Solution	Composition	
TAE 50 × buffer	50 mM	EDTA
	5.71 % (v/v)	acetic acid
	2 M	Tris-HCl
TE buffer	10 mM	Tris-HCl, pH 7.5
	1 mM	EDTA
Western Blot Blotting Buffer 5 ×	14.55 g	Tris
	7.33 g	glycine
	9.375 ml	10 % SDS
	dissolve in 1l ddH_2O; add 20 % methanol to 1 × Western Blot Blotting Buffer	
Media for mammalian cell culture		
SH-SY5Y medium	166 ml	DMEM
	30 ml	10 % FCS
	2 ml	P/S
	2 ml	L-glutamine
HeLa and HEK293-T medium	500 ml	DMEM
	55 ml	10 % FCS
	6 ml	P/S
	7 ml	L-glutamine
Bacterial growth media		
LB medium	5 g	yeast extract
	10 g	tryptone
	10 g	NaCl
	after solving in 800 ml ddH_2O, add ddH_2O to a final volume of 1l	
2 ×-YTA medium	10 g	yeast extract
	16 g	tryptone
	5 g	NaCl
	after solving in 800 ml ddH_2O, add ddH_2O to a final volume of 1l	

3.2.6 *E. coli* strains

Strain	Properties	Supplier
BL21	B F^- *dcm ompT hsdS*(r_B^- m_B^-) *gal* [mal_B^+]$_{K\text{-}12}$(λ^S)	Stratagene
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(StrR) *end*A1 *nup*G	Invitrogen

Strain	Properties	Supplier
XL10-Gold	TetR Δ(*mcrA*)183 Δ(*mcr*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E44 thi-1 *rec*A1 *gyr*A96 *rel*A1 *lac* Hte [F' *pro*AB *lac*I^{q}Z ΔM15 Tn10 (TetR) Amy CamR]	Stratagene

3.2.7 Human derived cell lines

Cell line	Tissue	Disease	Reference
HeLa	cervix	adenocarcinoma	[Gey *et al.*, 1952; Scherer and Hoogasian, 1954]
HEK293-T	kidney	–	[Graham *et al.*, 1977]
SH-S5Y5	bone marrow	neuroblastoma	[Biedler *et al.*, 1978]
U373	brain	glioblastoma astrocytoma	[Ponten and Macintyre, 1968]

3.2.8 Total RNA from human brain

Developmental stage	**Brain tissues** (all total RNA products were purchased from BioChain)
adult	total brain, cerebral cortex, frontal lobe, hippocampus, parietal lobe, temporal lobe
fetal	total brain, frontal lobe, parietal lobe, temporal lobe

3.2.9 Enzymes and restriction endonucleases

Miscellaneous enzymes	
Name	**Supplier**
Pfu Ultra High-Fidelity DNA Polymerase	Agilent Technologies
RQ1 RNase-Free DNase	Promega
SuperScript® III Reverse Transcriptase	Invitrogen
T4 RNA Polymerase	Promega
Taq DNA Polymerase	Qiagen

Restriction endonucleases	
Name	**Supplier**
*Bam*HI	Promega
*Dpn*I	New England Biolabs
*Eco*RI	Promega
*Hind*III	Fermentas
*Nde*I	Pharmacia Biotech
*Xho*I	New England Biolabs

3.2.10 Plasmids

Plasmid	Properties	Reference
pEYFP-N1	EYFP protein tag (C-terminal fusion), Km^{r}, Nm^{r}	Clontech
pmCherry-C1	mCherry protein tag (N-terminal fusion), Km^{r}, Nm^{r}	Clontech
pGEX-6P-3	GST protein tag (N-terminal fusion), PreScission™ protease cleavage site, Amp^{r}	GE Healthcare
pOTB7-*SARS*	*SARS* cDNA (IRAUp969H0860D) , Cm^{r}	imaGenes

3.2.11 Antibodies

Primary antibodies	Supplier
SARS, clone 1H4 ; monoclonal; raised in mouse	Abnova
GST; polyclonal; raised in goat	abcam
Lamin A/C; raised in mouse	Santa Cruz Biotechnology
Tubulin; raised in rat	abcam
Anti-GFP HRP; monoclonal; raised in mouse	abcam
Secondary antibodies	**Supplier**
bovine anti-goat IgG-HRP	Santa Cruz Biotechnology
goat anti-mouse IgG-HRP	Santa Cruz Biotechnology
mouse/human anti-rat IgG-HRP	Santa Cruz Biotechnology
Alexa 488-labeled goat anti-mouse IgG	Invitrogen

3.2.12 Primers

Primers for Family M289

Listed are the primers used for amplification and sequencing of the coding exons of *SARS* and *ZNF697*, RT-PCR as well as amplification and sequencing of *SARS* and *ZNF697* cDNA. Primers used for site-directed mutagenesis and cloning can be found in section 3.5.16. Primers used to sequence the other protein coding genes in interval chr1:107710128-120694597 are listed in the appendix (see section 9.5).

Primer	Sequence	Length
SARS_fw_ex1	TCACAGGCTGAGTGCTGC	268 bp
SARS_rv_ex1	GGTGGCTGGATTTGAGAGAG	
SARS_fw_ex2	GAGAAAGGAGTATCTTAATTATGGGC	193 bp
SARS_rv_ex2	GAGCAAGATCAAATATCTCCGC	
SARS_fw_ex3	TTATAAAACAATGCCAGAGATTTTC	221 bp
SARS_rv_ex3	TGAGAGCACTTGCTGTCTAGG	
SARS_fw_ex4	CTGTGCTGTCCTTGCCTCAC	298 bp
SARS_rv_ex4	TGTTCCACAGATCTCCTCCC	
SARS_fw_ex5	TTAGCCGAGCATAGTGTTGC	560 bp
SARS_rv_ex5	CCTGATAAGCCCTCTCTTGC	
SARS_fw_ex6	TTCTAGCTGAACATGATGGTGAC	295 bp
SARS_rv_ex6	ACAGCACCACTCTGCATCAC	
SARS_fw_ex7	CTAGCCAAGGTCCTCCCC	362 bp
SARS_rv_ex7	GCCTCAGCAACACAGATCC	
SARS_fw_ex8	GCAAGGATGTCTCCCACTTC	267 bp
SARS_rv_ex8	CTGGTAAGGGTGGCGTCTG	
SARS_fw_ex9	TTCTTCAGGCTTGCTAAAGG	298 bp
SARS_rv_ex9	AGAAATAATTCGGACGGGC	

Primer	Sequence	Length
SARS_fw_ex10	CCGTCAGTAAGACCCGATG	270 bp
SARS_rv_ex10	TATCCTGAAAGGGGACTGC	
SARS_fw_ex11	TGAGTCAGGACTGAGTTCTTTTAGC	252 bp
SARS_rv_ex11	CCTTTTCACCCCTTTCAAGC	
SARS_fw_ex12	GGTCTTAGGGCTTTGACTCAC	297 bp
SARS_rv_ex12	GGCTTCCCTGCTGTTGG	
ZNF697_fw_ex2	GTGGTGAGGGGATGAGACC	413 bp
ZNF697_rv_ex2	CCTTACAGAGAATGCTGTCGC	
ZNF697_fw_ex3.1	TTACTGAGGTTTCCTCTGGGC	411 bp
ZNF697_rv_ex3.1	GCGTCCATGATGCTGGC	
ZNF697_fw_ex3.2	GGCACCGAGGTGACAAG	381 bp
ZNF697_rv_ex3.2	GGTTGGTCAGGTAGGTGTTG	
ZNF697_fw_ex3.3	GAGCTGGATAGCCTGGTGG	313 bp
ZNF697_rv_ex3.3	GGTGGTTGGTCAGGTAGGTG	
ZNF697_fw_ex3.4	GATGGTGGGCATGATGG	307 bp
ZNF697_rv_ex3.4	GTGGCTCAACAGATGCGAG	
ZNF697_fw_ex3.5	CTTCAGCCGCAACACCTAC	395 bp
ZNF697_rv_ex3.5	GCTGGTGCTTCACCAAGTC	
ZNF697_fw_ex3.6	GCTTCGTGCGCCGTT	372 bp
ZNF697_rv_ex3.6	GCTGGTGCTGCGTGAG	
ZNF697_fw_ex3.7	GCACCTCTTCACGCACAAG	497 bp
ZNF697_rv_ex3.7	AAGGCTCCCCAGTCACTCTC	
Confirmation of cDNA synthesis		
HUWE1_ex_67	CAAGTGAGGAAAAGGGCAAA	569 bp
HUWE1_ex_68	GTTCATGAGCTGCCCCAGT	

Amplification and sequencing the cDNA of *SARS* and *ZNF697*

Primer	Sequence
cDNA-SARS-1*	GAAGATGGTGCTGGATCTGG
cDNA-SARS-2	GGTCACAGGCTGAGTGCTG
cDNA-SARS-3	GAGAGGATTTGGGGTGATTG
cDNA-SARS-4*	CTAAGCCCTCCCCAGAGATG
cDNA-SARS-5	TCGCCTTCAAAGCCATCTAC
cDNA-SARS-6	CTGATTGCCACCTCAGAGC
cDNA-SARS-7	TGCCCTCATGTTGCTTCTTC
cDNA-ZNF697-1*	AAGGGATTCCTGGTCACCTC
cDNA-ZNF697-2*	GTGCGTTTTATAGCGGAAGC
cDNA-ZNF697-3	ATGCTGTCCAAGGAACCATC
cDNA-ZNF697-4	CCGGGATATGCTGTCAGACT
cDNA-ZNF697-5	CAGATGTCGGGCACTGCTTC
cDNA-ZNF697-6	GGAGCTGGATAGCCTGGTG

Primer	Sequence
cDNA-ZNF697-7	AAGCGCTTCTCGCACTGG
cDNA-ZNF697-8	CTCGGGGAGCTGGATAGC
cDNA-ZNF697-9	GCCAGCTGAAGCTCTTGC

* primers were used for amplification of cDNA in RT-PCR

Primers for Family M8600485

Primer	Sequence	Length
ALDH5A1_fw_ex1	GCCTCCTCGCTCCTCTTG	531 bp
ALDH5A1_rv_ex1	GCTGGTGTCACTTTGGGG	
ALDH5A1_fw_ex2	AGCATTCTGTCTTACACTTGGC	223 bp
ALDH5A1_rv_ex2	GAAGTCAGCGGCTTTCCC	
ALDH5A1_fw_ex3	TTAGGAACACAGAGCCATGC	303 bp
ALDH5A1_rv_ex3	TGAATTGTTTCCCAACTCCC	
ALDH5A1_fw_ex4	TTGCACTAAGGAGGTGGTCC	256 bp
ALDH5A1_rv_ex4	CAAAATTGCTTCTTTCTGTCCC	
ALDH5A1_fw_ex5	GTGCACCCATTGTTTCCTG	400 bp
ALDH5A1_rv_ex5	GGGGGAGCTACTACATCAAGG	
ALDH5A1_fw_ex6	AGTCTGTCCCCAGTGTCAGC	377 bp
ALDH5A1_rv_ex6	CCAAAATTGGTGATCAGGATG	
ALDH5A1_fw_ex7	TCCCATGTACACCACTGTGC	408 bp
ALDH5A1_rv_ex7	AGGCAGTAGAGGTGGTGGG	
ALDH5A1_fw_ex8	TTTCACAGAGAGGCGGTAGC	346 bp
ALDH5A1_rv_ex8	CTATCCCCACCCCTCCAG	
ALDH5A1_fw_ex9	CCAGCATGCTTTATTGTTAACCTAC	418 bp
ALDH5A1_rv_ex9	CAACACAACTGTCTGCCTCC	
ALDH5A1_fw_ex10	CTGGTTTCCTTTCCTCTCCC	177 bp
ALDH5A1_rv_ex10	AAATCTGGCATGAGCTGGAG	
ALDH5A1_fw_ex11	AAGTCATCAATGGTGCCCTC	365 bp
ALDH5A1_rv_ex11	AAAATAATGGATGGCATGTACC	

3.3 Web resources

Service	URL	Reference
	Databases	
1000 Genome Project	`http://www.1000genomes.org/`	The 1000 Genomes Project Consortium [Abecasis *et al.*, 2010]
Exome Variant Server	`http://evs.gs.washington.edu/EVS`	NHLBI GO Exome Sequencing Project (ESP), Seattle, WA

Service	URL	Reference
HomoloGene	http://www.ncbi.nlm.nih.gov/homologene	[Coordinators, 2013]
OMIM ®	http://www.ncbi.nlm.nih.gov/omim	[McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD)]
Pfam	http://pfam.sanger.ac.uk/	[Finn *et al.*, 2010]
UCSC Genome Browser	http://genome.ucsc.edu/	[Kent *et al.*, 2002]
UniProt	http://www.uniprot.org/	[UniProt-Consortium., 2011; Jain *et al.*, 2009]
	Online tools	
Interpro	http://www.ebi.ac.uk/interpro/	[Quevillon *et al.*, 2005]
Primer3	http://frodo.wi.mit.edu/primer3/	[Rozen and Skaletsky, 2000]
UCSC Table Browser	http://genome.ucsc.edu/cgi-bin/hgTables?org=Human&db=hg19&hgsid=194023751&hgta_doMainPage=1	[Karolchik *et al.*, 2004]
	Pathogenetic or not predictions	
Mutation Taster	http://www.mutationtaster.org/	[Schwarz *et al.*, 2010]
Panther	http://www.pantherdb.org/tools/csnpScoreForm.jsp	[Thomas *et al.*, 2006, 2003]
PolyPhen2	http://genetics.bwh.harvard.edu/pph2/	[Adzhubei *et al.*, 2010]
PROVEAN	http://provean.jcvi.org/index.php	[Choi *et al.*, 2012]
SIFT	http://sift.jcvi.org/www/SIFT_seq_submit2.html	[Kumar *et al.*, 2009; Ng and Henikoff, 2006, 2003, 2002, 2001]
	Predictions of protein stability	
I-Mutant 2.0	http://folding.uib.es/i-mutant/i-mutant2.0.html	[Capriotti *et al.*, 2005a,b]
SCide	http://www.enzim.hu/scide/ide2.html	[Dosztanyi *et al.*, 2003]
MUpro	http://www.ics.uci.edu/~baldig/mutation.html	[Cheng *et al.*, 2006]
PoPMuSiC	http://babylone.ulb.ac.be/popmusic/	[Dehouck *et al.*, 2005]

3.4 Software

Programme	Supplier	Method
CodonCode Aligner	CodonCode Corporation	sequence assembly
Cyrillic 2	CyrillicSoftware	pedigree drawing
E.A.S.Y® Win32	Herolab	gel documentation
ImageQuant (Version 5.2)	Molecular Dynamics	quantitative analysis of images
ZEN 2009 Light Edition	Carl Zeiss MicroImaging GmbH	digital imaging

3.5 Molecular biology methods

3.5.1 Production of $CaCl_2$ competent cells

An overnight culture (100 µl glycerol stock in 5 ml LB medium) was incubated at 37 °C with vigorous shaking. The next day, 500 µl of the overnight culture was used to inoculate 30 ml LB medium. The culture was grown at 37 °C to the logarithmic phase of growth, OD_{600} 0.3, with vigorous agitation. Subsequently, 20 ml culture was centrifuged[1]. The supernatant was discarded and the pellet was dissolved in 5 ml of icecold 100 mM $MgCl_2$, gently mixed, and subsequently centrifuged. The supernatant was discarded and the pellet dissolved in 5 ml of icecold 100 mM $MgCl_2$ and gently mixed. After incubation on ice for 2 minutes the suspension was centrifuged again. After removal of the supernatant the pellet was dissolved in 1.5 ml of cold 100 mM $MgCl_2$ and mixed gently.

Aliquots of $CaCl_2$ competent cells in 85 % glycerol (100 µl) were stored at −80 °C. To determine the efficiency of the $CaCl_2$ competent cells, 50 µl freshly prepared cells were transformed with 100 ng control plasmid and plated on selective plates, containing the appropriate antibiotic.

3.5.2 Transformation

An aliquot of $CaCl_2$ competent cells was thawed on ice. After addition of 5 µl ligation mix or 100 ng plasmid DNA, cells were gently mixed and subsequently kept on ice for 60 minutes. An aliquot of $CaCl_2$ competent cells without addition of DNA was used as negative control.

The cells were then exposed to 42 °C for 45 seconds ("heat shock") and then chilled on ice for one minute. Subsequently, 1 ml of LB medium was added and cells were incubated at 37 °C for one hour with gentle agitation. Cells (100 µl, 150 µl and 100 µl, 150 µl in 1:10 dilution) were plated onto LB selection plates and incubated overnight at 37 °C.

3.5.3 Glycerol Stock Preparation

A single colony of a clone was picked from a plate and incubated overnight in selective LB medium. A 1 ml aliqout of the overnight culture was mixed with 1 ml of 85 % sterile glycerol in a sterile screw cap tube. The glycerol stock was stored at −80 °C.

3.5.4 Small scale DNA preparations – Miniprep

Small scale DNA preparations were carried out to analyse if clones contained the appropriate plasmid. Lysis under alkaline conditions was preferred if many clones had to be analysed.

Lysis under alkaline conditions

For isolation of plasmid DNA, individual colonies were inoculated in 10 ml LB medium containing the appropriate antibiotic. Cells were cultured at 37 °C overnight with vigorous agita-

[1] All centrifugation steps were performed at 5000 rpm, using rotor SS34, Sorvall RC-5 or 5B Refrigerated Superspeed Centrifuge (Dupont Instruments-Sorvall) at 4 °C for five minutes.

tion. The next day, 4 ml to 6 ml of cells were centrifuged at 14 000 rpm for one minute. Cells were resuspended in 200 µl miniprep buffer 1 and incubated for five minutes at RT. After addition of 400 µl miniprep buffer 2, reaction vials were inverted ten times, directly placed on ice and kept on ice for 30 minutes after addition of 300 µl miniprep buffer 3. Following addition of 400 µl chloroform and vortexing for ten seconds, the samples were centrifuged at 4 °C and 14 000 rpm for two minutes. The supernatant was transferred to a new 1.5 ml Eppendorf tube and 5 µl RNase (10 mg/ml) was added. After vortexing, samples were incubated at RT for ten minutes. Subsequently, 400 µl isopropanol was added, samples were vortexed and centrifuged at 4 °C and 14 000 rpm for 30 minutes. The supernatant was discarded and pellets were washed by addition of 500 µl 70 % ethanol and centrifugation at 4 °C, 14 000 rpm for 20 minutes. The supernatant was discarded and pellets were dried at RT. The pellet was dissolved in 50 µl EB buffer. Plasmid DNA was stored at −20 °C. All centrifugation steps were performed with centrifuge 5417R (Eppendorf).

Plasmid preparation using QIAprep® Spin Miniprep Kit

For the isolation of small amounts of plasmid DNA the QIAprep® Spin Miniprep Kit was used according to the manufacturer's instructions. Neutralized and cleared bacterial lysate was applied onto the spin column. DNA bound to the silicagel membrane was purified by serveral washing steps. Elution was performed with 30-50 µl EB buffer.

3.5.5 Large scale DNA preparations – Maxiprep

QIAfilter™Plasmid Maxi Kit

To obtain plasmid DNA yields of up to 500 µg maxi preps were performed with the QIAfilter™ Plasmid Maxi Kit according to the instructions of the manufacturer. Neutralized bacterial lysates were incubated in the QIAfilter™ Maxi Cartridge for ten minutes, purified by filtration and directly loaded onto the anion-exchange resin. After washing, plasmids were eluted from the resin and precipitated with isopropanol. Plasmid DNA pellets were dissolved in TE buffer.

Preparation of plasmid DNA with JETSTAR 2.0 ENDOTOXIN-FREE Plasmid Maxi Kit

Large scale endotoxin-free plasmid DNA preparations from *E. coli* TOP10 (see subsection 3.2.6) for transient transfection of mammalian cells were performed using JETSTAR 2.0 Plasmid Maxi Kit. This method allows to remove the last traces of bacterial endotoxin from plasmid DNA preparations.

The manufacturer's instructions were followed, with the exception of centrifugation time spans, which were doubled. Briefly, neutralized bacterial lysates were centrifuged at RT. The cleared lysate was loaded onto the equilibrated resin and allowed to enter the resin by gravity flow. The column was first washed with buffer ENDO-2, the second washing step was performed with standard wash buffer E5. After elution with buffer E6, isopropanol precipitation was car-

ried out. The DNA pellet was dissolved in 250 µl TE buffer at RT for 20 minutes. Plasmid DNA was stored at −20 °C. Dissolving was repeated twice.

3.5.6 DNA restriction digest

DNA restriction digests were carried out using conditions proposed by the restriction endonucleases' manufacturers. If restriction digest was carried out with two restriction endonucleases simultaneously, buffers were chosen in which both enzymes had at least 75-100 % activity. A general restriction digest contained 1 µg DNA, 2 µl of the respective restriction buffer, 0.2 µl BSA and 1 U/µg restriction endonuclease. Aqua ad iniectabilia was added to a final volume of 20 µl.

3.5.7 Ethanol precipitation of DNA

For precipitation of PCR fragments or plasmid DNA 2.5 volume of 100 % ethanol and 0.1 volume of 3 M sodium acetate, pH 5.2, were added to the sample. After incubation at −20 °C for 30 minutes or overnight, the sample was centrifuged at 14 000 rpm for 30 minutes. The supernatant was discarded carefully and the DNA pellet was washed with 500 µl 70 % ethanol. The sample was centrifuged at 14 000 rpm for 40 minutes. The supernatant was discarded completely. After drying at RT the DNA pellet was dissolved in 10–20 µl TE buffer. All centrifugation steps were performed with a 5417R centrifuge (Eppendorf).

3.5.8 Ligation

T4 DNA ligase catalyses the formation of a phosphodiester bond between the 5'-phosphate and the 3'-hydroxyl groups of adjacent nucleotides in double-stranded DNA. Ligation of plasmid DNA and insert was performed at 15 °C overnight. The amount of insert applied was five times higher than the amount of vector. As a standard condition 25 ng vector DNA was used. The amount of insert was calculated as shown in Equation 3.1. The ligation mixture was composed of 1 µl T4 ligase, 2 µl 10 × buffer, 25 ng Plasmid DNA, an appropriate amount of PCR fragment and ddH_2O in a final volume of 20 µl.

$$\text{concentration}_{\text{insert}}[\text{ng}] = 125[\text{ng}] \cdot \frac{\text{length}_{\text{insert}}[\text{bp}]}{\text{length}_{\text{vector}}[\text{bp}]} \tag{3.1}$$

3.5.9 Agarose gel electrophoresis

Depending on DNA fragment size, 0.7–2 % (w/v) agarose gels were prepared. Agarose was boiled in TAE buffer in a microwave. After a short cooling down period, ethidium bromide was added (1 µl per mg agarose), and the still liquid gel-solution was poured into an electrophoresis gel tray. An appropriate comb was used to produce gel pockets. After polymerisation, the comb was removed and the samples were pipetted into the gel pockets. Agarose gel electrophoresis was carried out at 150 mV in an electrophoresis chamber containing TAE buffer. Gels were documented using the E.A.S.Y Win32® (Herolab) system.

3.5.10 Polymerase chain reation – PCR

PCR fragments that were used for cloning or *in vitro* transcription were amplified using *Pfu* polymerase. Different PCR reaction mixtures (Table 3.3) were used depending on template quality and subsequent procedures. Elongation time depended on the size of the amplicon and the polymerase used: *Taq* polymerase amplifies 1 kb in 30 seconds, *Pfu* polymerase amplifies 1 kb per minute. DNA was denatured at 96 °C. The elongation step of PCR was performed at 72 °C, with the exception of FailSafe™ reactions, which required an elongation temperature of 68 °C. In general, touch-down PCR with a stepwise reduction of annealing temperature (0.5 °C per cycle), was applied first for most PCR reactions.

Table 3.3: PCR mixtures and cycling conditions

***PfuUltra* Polymerase Mix**		**FailSafe™**	
Aqua ad iniectabilia	40.6 µl	DNA template	1 µl
10 × *PfuUltra* HF reaction buffer	5 µl	Forward primer	3 µl
dNTPs (25 mM each dNTP)	0.4 µl	Reverse primer	3 µl
DNA template	1 µl	FailSafe™	10 µl
Forward primer	1 µl	Taq Polymerase	0.3 µl
Reverse primer	1 µl	Aqua ad iniectabilia	3 µl
PfuUltra HF (2.5 U/µl)	1 µl		
Bio-X-Act™ Long Mix		**Mager Mix**	
DNA template	1 µl	DNA template	1 µl
Forward primer	1 µl	Forward primer	1 µl
Reverse primer	1 µl	Reverse primer	1 µl
Bio-X-Act™Long Mix	12.5 µl	Mager Mix	45 µl
Aqua ad iniectabilia	9.5 µl		

Primers were used at a concentration of 10 µM and template DNA was used at concentrations of 15–30 ng/µl, respectively. PCRs which failed to produce amplicons with Mager Mix conditions were conducted either with Bio-X-Act™ Long Mix or FailSafe™.

3.5.11 PCR reaction for Sanger sequencing

PCR fragments and plasmids were sequenced according to Sanger's Dideoxy DNA sequencing method, where a different fluorescent dye is attached to each of the four ddNTPs. If necessary, PCR fragments were purified using the QIAquick Gel Extraction Kit and plasmid DNA was prepared using a plasmid preparation kit before the sequencing reaction. The sequencing reaction was set up with primers used for PCR amplification or primers binding within the PCR amplicon. In the case of homopolymeric adenine or thymine regions (e. g. STXBP3 exon 18), anchored primers were used to avoid polymerase slippage. Betaine was used in the event of problematic sequencing reactions, e. g. with GC-rich templates. In that case, PCR product, betaine and water were incubated at 96 °C for two minutes before the addition of primer, BigDye® and BigDye® Terminator. Different protocols were used for sequencing PCR products or plasmid DNA (Table 3.4).

Table 3.4: PCR composition and cycling parameters for Sanger Sequencing

PCR product				
DNA	2 ng/100 bp	96 °C	1 min	
BigDye	2 µl	96 °C	30 sec	
BigDye Terminator	2 µl	50 °C	15 sec	25 cycles
Primer 10 µM	1 µl	60 °C	4 min	
Aqua ad iniectabilia	add to 10 µl	4 °C	∞	
Plasmid				
DNA	80 ng	96 °C	1 min	
BigDye	2 µl	96 °C	10 sec	
BigDye Terminator	2 µl	50 °C	5 sec	25 cycles
Primer 10 µM	2 µl	60 °C	4 min	
Aqua ad iniectabilia	add to 10 µl	4 °C	∞	

Precipitation of labeling reaction

Prior to electrophoresis and base calling, which was conducted by the in-house service facility, precipitation of the labeling reaction-product was performed. After addition of 1 µl 2 % SDS, samples were incubated at 98 °C for ten minutes. Subsequently, 25 µl 100 % ethanol was added and the reaction was mixed thoroughly by inverting. Samples were centrifuged at 4 °C and 4000 rpm for 60 minutes. The supernatant was discarded carefully and 150 µl 70 % ethanol was added for washing the DNA pellet. Samples were centrifuged at 4 °C and 4000 rpm for 30 minutes. The washing step was repeated once. The pellet was dried by adjusting the plate upside-down on a paper towel and centrifuging shortly up to 4000 rpm. All centrifugation steps were performed with Centrifuge 5810R (Eppendorf). Electrophoresis and base calling was carried out with ABI 3130xl 16-capillary and 3730xl 96-capillary DNA analyzer. Sequencing results were analysed with the CodonCode Aligner software (CodonCode Corporation).

3.5.12 RNA extraction from cell lines

To inhibit RNase activity, 10 ml TRIzol® reagent was added to a 50 ml cell pellet (5×10^7 cells). After incubation at RT for 30 minutes, the tube was shaken until a homogeneous solution was obtained. The suspension was transferred to a 30 ml RNase free tube. After addition of 2 ml chloroform the suspension was mixed vigorously for 15 seconds, was then kept at RT for three minutes and was subsequently centrifuged[2]. The supernatant was transferred to a fresh 30 ml RNase free tube and mixed with 5 ml isopropanol. After incubation at RT for five minutes, the sample was centrifuged again. The supernatant was discarded and 10 ml of 70 % ethanol was

[2] All centrifugation steps used in this protocol were performed at 5000 rpm, using rotor SS34 in a Sorvall RC-5 centrifuge or 5B Refrigerated Superspeed centrifuge (Dupont Instruments-Sorvall) at 4 °C for 20 minutes, unless otherwise indicated.

added to the pellet and mixed by gentle shaking. Again, the sample was centrifuged for ten minutes and the supernatant discarded. The pellet was dried and RNA was dissolved in 600 ml DEPC-H_2O. The sample was kept on ice for ten minutes and was subsequently incubated at 65 °C for five minutes. RNA concentration was determined spectrophotometrically and the quality was checked on an agarose gel. RNA was stored at −20 °C.

3.5.13 Synthesis of cDNA

For cDNA synthesis either commercially available total RNA or total RNA from freshly prepared cell line extracts was used. The reaction mixture was prepared (total RNA 1 µg, random primers 2 µl, 10 mM dNTPs 1 µl, and aqua ad iniectabilia in a final volume of 21 µl) and incubated at 65 °C for five minutes. After incubation on ice for two minutes, 5 × first strand buffer (6 µl), 0.1 M DTT (1 µl), RNasin (1 µl), and SuperScript® III (1 µl) were added and after gentle mixing the sample was incubated at 20 °C for five minutes, then at 50 °C for 60 minutes and finally at 70 °C for 15 minutes.

3.5.14 Reverse transcription polymerase chain reaction – RT-PCR

To verify the success of cDNA synthesis, exon-spanning primers binding in *HUWE1* exons 67 and 68 (hg19; subsection 3.2.12) were used, resulting in an 569 bp amplicon. For amplification of the full *SARS* transcript, primers cDNA-SARS-1 and cDNA-SARS-4 (see subsection 3.2.12) were used, giving rise to a 1786 bp amplicon. *ZNF697* cDNA was amplified using primers cDNA-ZNF697-1 and cDNA-ZNF697-2 (see section 3.2.12) resulting in a 1645 bp amplicon. The RT-PCR conditions listed in table 3.5 were used, with primer concentrations of 10 µM and cDNA concentration of 30 ng/µl, respectively.

3.5.15 Site-directed mutagenesis

Site-specific mutagenesis of *SARS* and *ZNF697* was performed with the QuickChange™ II XL Site-Directed Mutagenesis Kit according to the manufacturer's instructions. Mutagenic primers were designed using the QuikChange Primer Design Program (`http://www.stratagene.com/qcprimerdesign`). Plasmids isolated via the QIAfilter™ Plasmid Maxi Kit were used as templates for site-directed mutagenesis reactions.

3.5.16 Cloning of pmCherry-C1, pEYFP-N1 and pGEX6P3 constructs

Cloning of *SARS* cDNA

A full-length *SARS* cDNA clone (IRAUp969H0860D) was purchased from ImaGenes GmbH. The cDNA was isolated from a neuroblastoma cell line and contained a change from cytosine to thymine at c.1530, resulting in a missense mutation from arginine to cysteine at position 435 in the SARS protein. To correct this c.1530C > T point mutation, two *SARS* cDNA segments were amplified using primers containing the wild-type nucleotide cytosine as well as restriction sites for *Nde*I and *Bam*HI or *Nde*I and *Xho*I (Figure 3.3). Both PCR products were restricted with

Table 3.5: RT-PCR conditions for the use of *SARS*, *ZNF697* and *HUWE1* primers

SARS		96 °C	3 min	
cDNA	1 µl	94 °C	30 sec	
cDNA-SARS-1	1 µl	58 °C	30 sec	32 cycles
cDNA-SARS-4	1 µl	72 °C	45 sec	
Bio-X-Act™ Long Mix	12.5 µl	72 °C	10 min	
Aqua ad iniectabilia	9.5 µl	15 °C	∞	
ZNF697		96 °C	3 min	
cDNA	1 µl	94 °C	15 sec	
cDNA-ZNF697-1	3 µl	55 °C	45 sec	30 cycles
cDNA-ZNF697-2	3 µl	68 °C	4 min	
FailSafe™	10 µl	68 °C	10 min	
Taq DNA Polymerase	0.4 µl	15 °C	∞	
Aqua ad iniectabilia	3 µl			
HUWE1				
cDNA	1 µl	94 °C	3 min	
Bio-X-Act™ Long Mix	12.5 µl	94 °C	30 sec	
HUWE1_ex_67	1 µl	55 °C	45 sec	35 cycles
HUWE1_ex_68	1 µl	72 °C	45 sec	
Aqua ad iniectabilia	9 µl	15 °C	∞	

*Nde*I and ligated to obtain full-length *SARS* cDNA. The ligation product was amplified by PCR using primers SARS_TS1_F and SARS_TS2_R (Table 3.6). Subsequently, wild-type *SARS* cDNA was cloned into pmCherry-C1 using *Bam*HI and *Xho*I restriction sites.

To obtain *SARS* c.514G > A (p.D172) and c.1285A > G (p.T429A), site-directed mutagenesis was performed as described using sense and antisense primers carrying the respective nucleotide changes. The pmCherry-C1-*SARS* constructs were used as templates in PCRs for cloning SARS wild-type and mutant cDNA into pEYFP-N1 and pGEX6P3 (Figures 3.5 and 3.4). The primers containing the appropriate restriction sites are listed in Table 3.6.

Table 3.6: *SARS* cloning primers

Removal of *SARS* c.1530C > T mismatch		
Primer	**Sequence**	**Length**
SARS_TS1_F	TGTCGCTCGAGCGATGGTGC	1298 bp
SARS_TS1_R	TAGCATTGAGCATATGGACAAACTCCACCT	
SARS_TS2_F	TGTCCATATGCTCAATGCTACCATGTGCGCCACTACCCGTACC	291 bp
SARS_TS2_R	GTGCCGGATCCTCAAGCATCGGTGACCTCCA	

Side directed mutagenesis of *SARS* cDNA		
Primer	**Sequence**	**Length**
Sense D172N	AGAAGTACTCTCATGTGAACCTGGTGGTGATGGTA	–
Antisense D172N	TACCATCACCACCAGGTTCACATGAGAGTACTTCT	
Sense T429A	GTCCATATGCTCAATGCTGCCATGTGCGCCACTA	–
Antisense T429A	TAGTGGCGCACATGGCAGCATTGAGCATATGGAC	
Cloning of *SARS* cDNA into pEYFP-N1		
Primer	**Sequence**	**Length**
SARS-XhoI-N	CATCGCTCGAGAAGAAGATGGTGCTGGATCTGG	1572 bp
SARS-BamHI-N	CGTGCGGATCCTCAGCATCGGTGACC	
Cloning of *SARS* cDNA into pGEX6P3		
Primer	**Sequence**	**Length**
pGEX-6P-3-F	TATATGGATCCATGGTGCTGGATCTG	1567 bp
pGEX-6P-3-R	CTATACTCGAGTCAAGCATCGGTGAC	

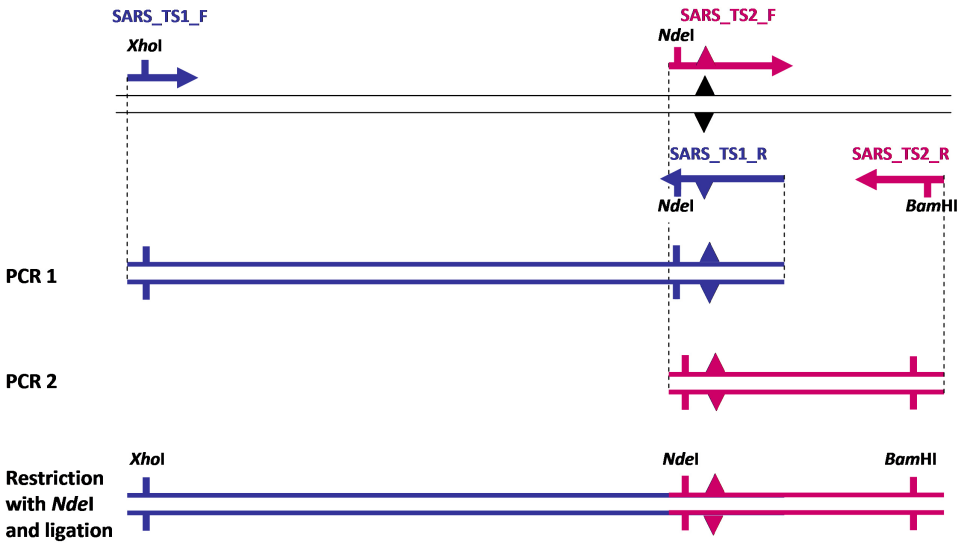

Figure 3.3: PCR-strategy for removing *SARS* c.1530C > T point mutation from cDNA clone template. Primers are depicted as arrows, respective restriction sites are depicted as short vertical lines, the position of the base substitution of the wild-type allele cyctosine with thymine (c.1530C > T) is depicted as triangle.

Cloning of *ZNF697* cDNA

Full-length cDNA of *ZNF697* was cloned from human adult brain mRNA (BioChain) via RT-PCR. The primers used for amplification are listed in table 3.7. The cDNA was inserted via *Bam*HI and *Hind*III restriction sites into pmCherry-C1, and the resulting construct was used as template in site-directed mutagenesis introducing c.472C > A (p.P158T) using the appropriate

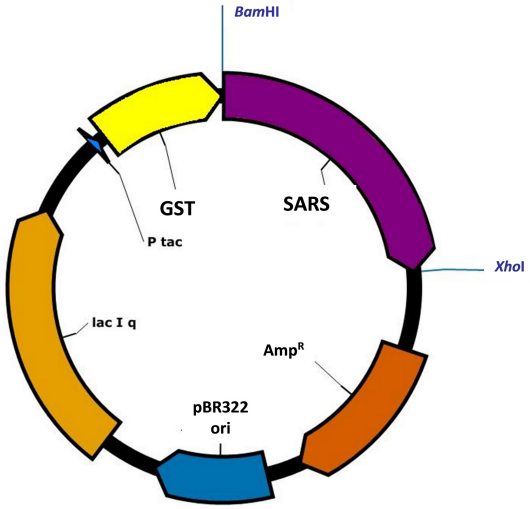

FIGURE 3.4: SARS-pGEX6P3 construct. P_{tac}: tac promoter, GST: glutathione S transferase, Amp^R: ampicillin resistance gene, pBR322 ori: origin of replication, *lacIq*: *lacIq* repressor gene; the restriction sites for *Bam*HI and *Xho*I are indicated.

sense and antisense primers (Table 3.7). Subsequently, the pmCherry-C1-*ZNF697* constructs were used as templates in PCRs for cloning *ZNF697* cDNA into pEYFP-N1 (Figure 3.5).

Table 3.7: *ZNF697* cloning primers

Side directed mutagenesis of *ZNF697* cDNA		
Primer	**Sequence**	**Length**
Sense P158T	CGAGGTGACAAGACCGCCCACCGCC	–
Antisense P158T	GGCGGTGGGCGGTCTTGTCACCTCG	
Cloning of *ZNF697* cDNA into pmCherry-C1 (C) and pEYFP-N1 (N)		
Primer	**Sequence**	**Length**
ZNF697-HindIII-C	CATCGAAGCTTCCTGGATGAAACAAGAAG	1668 bp
ZNF697-BamHI-C	GCATAGGCGGATCCCTAACACAGGTGC	
ZNF697-HindIII-N	CATCGAAGCTTACCTGGATGAAACAAGAAGATAATC	1665 bp
ZNF697-BamHI-N	CGTGCGGATCCGTACACAGGTGC	

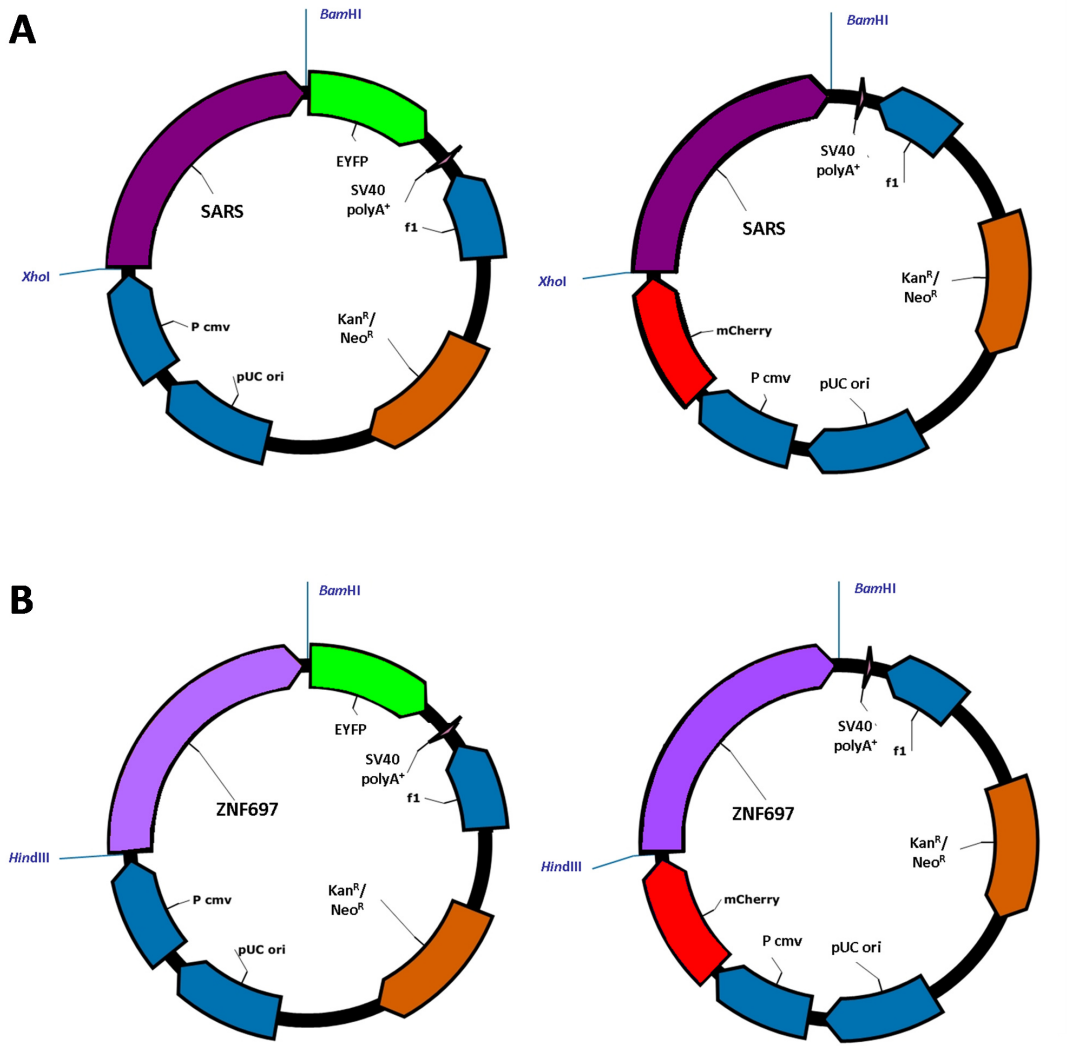

FIGURE 3.5: **A**) *SARS*- and **B**) *ZNF697*-constructs used for ectopic gene expression in mammalian cell lines. cDNAs were cloned into pEYFP-N1 and pmCherry-C1 expression vectors. P_{cmv}: human cytomegalovirus immediate early promoter, SV40 poly A+: SV40 early poly A+ signals, f1: f1 origin of replication, Kan^R/Neo^R: kanamycin/neomycin resistance gene, ori: origin of replication; the restriction sites for *Bam*HI, *Hind*III and *Xho*I are indicated.

3.6 Protein biochemical methods

3.6.1 Expression and purification of GST-fusion proteins

The pGEX6P3 vector was used for high-level inducible intracelluar expression of SARS wild-type and mutant proteins in *E. coli* BL21. A single colony containing a recombinant pGEX6P3 plasmid was inoculated into 100 ml 2× TYA medium containing 50 µg/ml ampicillin and incubated overnight at 37 °C with vigorous shaking. The following day, the overnight culture was diluted into 1 l of 2× TYA containing 50 µg/ml ampicillin and grown to an OD_{600} of 0.68 with vigorous agitation at 37 °C. Expression of GST-fusion protein was induced by addition of 1 mM IPTG. After incubation at 37 °C with vigorous agitation for four hours, liquid cultures were centrifuged at 4 °C at 7000-10 000 rpm for 30 minutes[3]. Pellets were stored at −80 °C.

Batch purification with Glutathione Sepharose 4B

Pellets were thawed slowly on ice, dissolved in ice-cold 1 × PBS containing an appropriate amount of complete protease inhibitor tablets and aliquots of 5 ml were then transferred into 15 ml tubes. Sonication was performed using an ultrasonic cell disruptor (Bandelin) at maximum power for ten seconds and two cycles. Afterwards sonicates were centrifuged to remove insoluble material (500 × g, 4 °C, 20 minutes).

Glutathione Sepharose 4B was prepared according to the manufactures instructions. 2 ml of 50 % Glutathione Sepharose 4B slurry was added to each 100 ml of lysate sample. The mixture was incubated at 4 °C on an overhead shaker overnight. The next day, the matrix was sedimented by centrifugation at 4 °C at 500 × g for ten minutes. The supernatant was decanted and stored at −80 °C for further analysis. The matrix was washed by addition of 5 ml binding buffer to each 1 ml of 50 % Glutathione Sepharose 4B slurry, incubated at 4 °C on an overhead shaker for five minutes and centrifuged at 4 °C at 500 × g for five minutes. The washing step was repeated three times. GST-fusion protein was eluted by addition of 0.5 ml elution buffer per 1 ml of Glutathione Sepharose 4B, incubation at RT on an overhead shaker for five minutes and sedimentation of the matrix via centrifugation at 4 °C at 500 × g for ten minutes. The protein-containing supernantant was removed carefully and kept on ice. Elution was repeated for five to six times and protein concentration was determined spectrophotometrically, and/or by Coomassie staining or by silver staining.

3.6.2 Buffer exchange

Buffer exchange was performed with PD-10 Desalting Columns (GE Healthcare). The protocol by gravity flow was applied according to the manufacturer's instructions.

[3] Rotors GSA or GS3, Sorvall RC-5 or 5B Refrigerated Superspeed Centrifuge (Dupont Instruments-Sorvall) were used.

3.6.3 SDS cell lysis

SDS cell lysis was carried out in small-scale experiments analysing the overexpression of GST-fusion proteins. Collection of cells from 5 ml cell suspension was performed via centrifugation at 13 000 rpm for one minute. The supernatant was discarded and pellets were resuspended in SDS lysis buffer (30 µl per 1 OD_{600}, e. g. if $OD_{600} = 10$, 300 µl was used). After incubation at 60 °C for five minutes, samples were incubated at 95 °C for five minutes. Subsequently, probes were centrifuged at RT and 13 000 rpm for 30 minutes (Centrifuge 5417R, Eppendorf). The supernatant was transferred to a new 1.5 ml Eppendorf tube. For SDS-PAGE samples were mixed with the appropriate amount of Magic Mix (see subsection 3.2.5) and incubated at 95 °C for five minutes prior to gel loading.

3.6.4 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out using Mini-PROTEAN™ 3 Electrophoresis System (Bio-Rad). SDS-gels were prepared according to table 3.8. Running gel mixture was prepared and covered with a layer of isopropanol to make the surface even. After polymerisation, isopropanol was removed and after washing with ddH_2O the stacking gel mixture was applied and the comb placed. After polymerisation the comb was removed and samples were applied. Prior to SDS-PAGE, samples were mixed with an appropriate amount of Magic Mix (see subsection 3.2.5) and incubated at 95 °C for five minutes. Electrophoresis was carried out first at 100 V until the samples had passed through the stacking gel, then 180 V were applied until the designated electrophoretic separation was attained. After SDS-PAGE, gels were either stained with Imperial™ Protein Stain (see subsection 3.6.5), silver (see subsection 3.6.6) or used for Western Blotting (see subsection 3.6.7).

Table 3.8: Composition of stacking and running gel for SDS-PAGE

Stacking Gel	
ddH_2O	4.0 ml
30 % acrylamide mix	3.3 ml
1.5 M Tris-HCl, pH 8.8	2.5 ml
10% SDS	0.1 ml
10% ammonium persulfate	0.1 ml
TEMED	0.004 ml

Running Gel	
ddH_2O	1.4 ml
30 % acrylamide mix	0.33 ml
1.0 M Tris-HCl, pH 6.8	0.25 ml
10% SDS	0.02 ml
10% ammonium persulfate	0.02 ml
TEMED	0.002 ml

3.6.5 Protein staining in SDS-polyacrylamide gels

For protein staining in SDS-polyacrylamide gels the Coomassie R-250 dye-based reagent Imperial™ Protein Stain (see subsection 3.6.5) was used according to the manufacturer's instructions. After the final washing step, gel was documented using the E.A.S.Y Win32® (Herolab) system. For determining the protein concentration in gel bands, the SDS-polyacrylamide gels were scanned and then analyzed using the ImageQuant software (Molecular Dynamics).

3.6.6 Silver staining of polyacrylamide gels

The ProteoSilver™ Plus Silver Stain Kit (Sigma) was used for detection of very low protein concentrations. In short, after electrophoresis the polyacrylamid gel was immersed in Fixing solution overnight. Afterwards, the gel was washed with 100 ml of 30 % ethanol for 10 minutes. Then, the gel was washed with 200 ml of ultrapure water for 10 minutes. Sensitization was carried out for 10 minutes using 100 ml of Sensitizer solution. After another wash with 200 ml of ultrapure water, Silver solution (100 ml) was applied and the gel was incubated for 10 minutes. A short washing step was conducted with 200 ml of ultrapure water for 60 seconds, then the gel was developed for 7–10 minutes in Developer solution. After application of ProteoSilver Stop solution the gel was incubated for five minutes, before washing with 200 ml of ultrapure water. For determining the protein concentration in gel bands, gels were scanned and analyzed using the ImageQuant software (Molecular Dynamics).

3.6.7 Semi-dry Western Blot

Before assembly of the blot apparatus, Whatman 3MM chromatography paper was pre-soaked with 1 × blotting buffer and PVDF Western Blotting Membrane was first dampened with methanol, then also pre-soaked with 1 × blotting buffer. The transfer stack was build from cathode to anode as follows: six layers Whatman 3MM chromatography paper, SDS-gel, PVDF Western Blotting Membrane, six layers Whatman 3MM chromatography paper. Proteins were transferred for 40–90 minutes, with 20 V.

3.6.8 Immunodetection of proteins

Following Western Blot, PVDF membranes were incubated in blocking buffer (5 % milk or 3 % BSA) at RT on a roll mixer for 30–60 minutes to block unspecific binding sites. Primary antibodies were diluted adequately in blocking buffer and primary antibody incubation was performed at 4 °C on a roll mixer overnight. The next day, membranes were washed at RT in 1× PBST three times for five minutes. Next, membranes were incubated with secondary antibodies diluted adequately in 1× PBST. Incubation was performed at RT on a roll mixer for one hour. Subsequently, membranes were washed in 1× PBST three times at RT for five minutes. To detect immunolabeled proteins, Western Lightning®-ECL, Enhanced Chemiluminescence Substrate was used. Oxidizing Reagent and Enhanced Luminol Reagent were mixed at a ratio of 1:3 and transferred onto the PVDF membranes. Membranes were placed between colour laser transparency film and air pockets were smoothed out. FUJI Medical X-Ray Films were exposed to the membrane and developed using a CURIX 60™ table-top processor.

3.6.9 Cell fractionation analysis

EYFP-SARS wild-type and mutant expression vectors were transfected into HEK293-T cells. After 24 hours cells were harvested by centrifugation at 1000 rpm for five minutes using a Rotanta

46K centrifuge (Hettich) and washed once in 1 × PBS. The pellet was resuspended in 400 µl Cell Fractionation Lysis Buffer and incubated at RT for 10 minutes, then centrifuged at 3000 rpm for 10 minutes (Centrifuge 5417R, Eppendorf). The supernatant, which contained the cytosolic fraction, was removed and the pellet was washed twice with Cell Fractionation Lysis Buffer. The pellet, which contained the membrane and nuclear faction, was resuspended in 150 µl Magic Mix (see subsection 3.2.5).

3.7 Mammalian cell culture

3.7.1 Passage of cells

Used medium was aspirated and cells were rinsed by addition of 5 ml 1 × PBS and gentle swinging of the cell culture flask. 1 × PBS was then aspirated and 5 ml Trypsin/EDTA solution (Lonza) was added. After incubation at 37 °C for three minutes, cells were detached by tapping the flask against a hard surface and success of trypsinization (confluency of cells) was checked via microscope. Then, cells were diluted in an appropriate volume of fresh growth medium.

3.7.2 Transfection of cell-culture cells

The day before transfection, 1.5×10^5 cells were seeded per well in a six-well plate containing sterile cover slips. Cells were incubated at 37 °C in a CO_2 incubator for 24 hours. At the time of transfection, cells were 90–95 % confluent. Mix I (1 µg DNA, 100 µl OPTI-MEM) was prepared and incubated at RT for five minutes. Mix I and Mix II (1.5 µl LipoFECTAMINE™2000 Reagent, 100 µl OPTI-MEM) were gently mixed and incubated at RT for 20 minutes. Next, medium was aspirated from the wells and 1.5 ml fresh, antibiotic-free medium was added. Subsequently, 500 µl of the transfection mix was added to the cells. After gentle mixing the cells were incubated in a CO_2 incubator for 24–48 hours.

3.7.3 Fixation of cells

After transfection for 24–48 hours, medium was aspirated from the wells and cells were washed shortly with 1 ml 1 × PBS. Next, 1 ml 4 % PFA solution was added to each well. After incubation at RT for ten minutes PFA solution was aspirated and cells were washed again shortly with 1 × PBS. For mounting, 10 µl DAPI 1:10 000 in Fluoromount ™ were applied onto a glass slide and subsequently coverslips with immobilised cells were transferred onto the glass slide.

For staining of endogenous proteins, cells were first washed with 1 × PBS following fixation. Then, cells were permeabilized either with methanol or Triton® X-100 prior to primary antibody incubation.

Cell permeabilization with methanol

For permeabilization with methanol, cells were covered completely with ice-cold 100 % methanol and then incubated for 10 minutes at −20 °C. Afterwards, methanol was aspirated and cells were washed three times with 1 ml 1 × PBS for five minutes, respectively. Subsequently, cells were blocked with 2 % BSA for 60 minutes at RT. Blocking solution was then aspirated and cells were incubated with the primary antibody overnight at 4 °C. The next day, cells were washed three times with 1 ml 1 × PBS for five minutes, respectively. Subsequently, cells were incubated in fluorochrome-conjugated secondary antibody for 60 minutes at RT in the dark. After washing three times with 1 ml 1 × PBS for five minutes, coverslips were attached onto glass sildes with fluoromount.

Cell permeabilization with Triton® X-100

Cells were incubated in 0.2% Triton® X-100 in 1 × PBS at RT for ten minutes. Afterwards, the solution was aspirated and cells were washed three times with 1 ml 1 × PBS for five minutes, respectively. Blocking and secondary antibody incubation was performed as described for methanol permeabilization.

3.8 Enzyme activity measurements

3.8.1 Pyrophosphate release assay

Pyrophosphate release of purified recombinant GST-SARS wild-type and mutant (p.D172N and p.T429A) proteins and the GST protein was assayed at 37 °C for 90 minutes. Per reaction, 12.5 ng of protein were used in a total volume of 300 µl. Negative controls contained buffer eluated from PD-10 desalting columns instead of protein.

Pyrophosphate release was determined at three different time points (0, 60 and 90 minutes). Reactions were stopped approximately simultaneously by adding twofold reaction volume of BIOMOL GREEN Reagent™ to each sample. Samples were then incubated at RT for 20 minutes to allow for the development of the green color. Subsequently, 300 µl of this mix per sample were each pipetted into a well of a 96 well EIA/RIA-Plate (Corning) and OD_{620} was measured in a microplate reader (BMG Labtech).

Standards were prepared in pyrophosphate release reaction buffer that had been pre-incubated for 90 minutes at 37 °C. After incubation, serial dilutions were prepared to obtain concentrations of 2, 1, 0.5, 0.25, 0.125, 0.63 and 0.31 nmol phosphate, respectively. Standard sample reactions were terminated at the same time as assay sample reactions using twofold reaction volume of BIOMOL GREEN Reagent™. Standard samples were prepared in triplicate for each phosphate concentration.

3.8.2 Measurement of SSADH activity

EBV immortalized lymphoblastoid cell lines derived from the blood of the index patient (V-11, family 8600485) as well as from six German age and sex matched controls were established using standard procedures. The method used is based on the method developed by Gibson and colleagues [Gibson *et al.*, 1991], which is in turn based on the fluorometric assay by Cash and co-workers [Cash *et al.*, 1977].

Cell extracts were prepared by sonication of cell pellets (4×10^7 cells/ml) in 100 mM Tris-HCl (pH 8.6). Cell extracts were sonicated twice for 2 cycles each of 10 seconds at maximum power with a pause between the bursts and cells were kept on ice during the whole procedure. Subsequently, the cell lysates were applied twice onto QIAshredder homogenizer columns (Qiagen) to reduce viscosity caused by cell debris and high-molecular-weight cellular components. Total protein concentration was measured spectrophotometrically. SSADH activity was assayed for 30 minutes at 37 °C using 100 µl cell lysate, 3 mM succinic semialdehyde (SSA), 0–2 mM NAD^+ and SSADH-buffer in a final volume of 300 µl. Blank reactions contained water substituted for SSA. The reaction was stopped by heating for five minutes at 100 °C. After cooling down on ice for 10 minutes, probes were centrifuged at 4 °C (20 800 g). The supernatant was removed and centrifuged again using the same conditions as before. Then, 250 µl of each sample were pipetted per well of a 96 Well FIA-Plate (Greiner Bio-One). NADH fluorescence was measured in a microplate reader (BMG Labtech; excitation 355 nm, emission 460 nm). Probes and standard samples were measured in triplicates. NADH standards were prepared as serial dilution (3.125–200 mM NADH) using 1 mM NADH stock solution.

4 Results – Family M289

4.1 Linkage analysis and mutation screening in family M289

Parametric linkage analysis in family M289 revealed a single interval of homozygosity on chromosome 1p13.3–p11.2 near the centromer with the highest probable significant LOD score above four (Figures 4.1)[1]. The linkage interval had a size of 12.9 Mb, was flanked by the heterozygous SNPs rs10494061 and rs1938250[2] and contained 133 RefSeq genes (hg19) (Figures 4.2 and 4.3). In this study, coding regions including intron-exon boundaries of these genes were amplified by PCR and subsequently sequenced by Sanger sequencing.

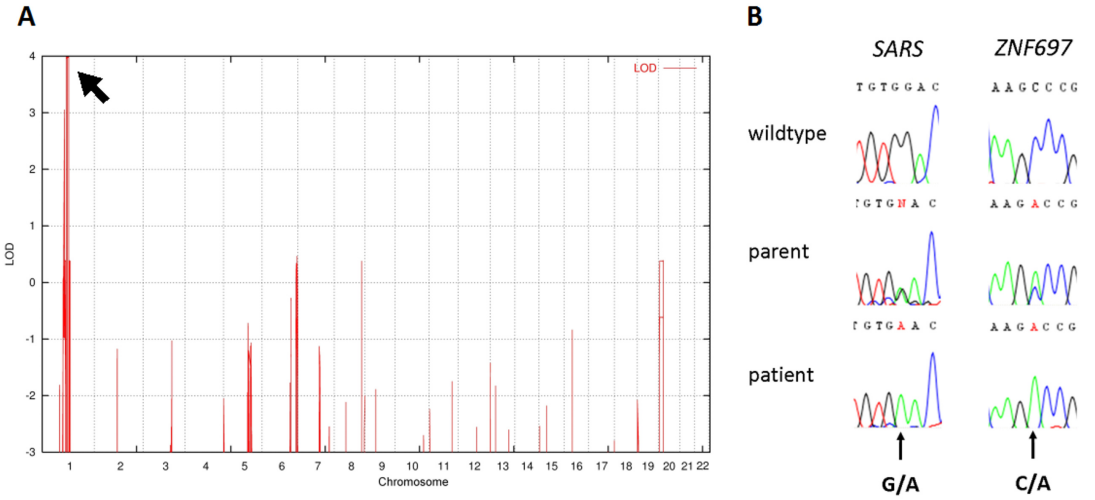

FIGURE 4.1: **A)** Linkage profile resulting from whole genome parametric linkage analysis based on the assumption of second cousin marriage, showing a single linkage interval (arrow) with significant LOD score > 4 on chr.1p13.3–p11.2. **B)** Sequence chromatograms of an affected individual, a parent and a healthy control (wild-type), showing homozygous substitutions c.514G > A in *SARS* and c.472C > A in *ZNF697* in the affected individual.

Two missense mutations co-segregating with the disease were detected in the interval (Figure 4.1 B and Figure 4.3). In *SARS* (seryl-tRNA synthetase; NM_006513) a G > A substitution (c.514G > A, chr1: 109773566; hg19) was found in exon five and results in a change from as-

[1] Linkage analysis was carried out by Dr. M. Garshasbi. Genotyping (SNP analysis) of all affected family members, their parents and healthy siblings was performed using the Human Mapping 50 K Array, Version 2 (Affymetrix) [KENNEDY *ET AL.*, 2003] based on previously published protocols [MATSUZAKI *ET AL.*, 2004]. Linkage analysis was performed using the Merlin software. Details of data quality controls and linkage analysis have been published elsewhere [GARSHASBI *ET AL.*, 2006].

[2] rs10494061: chr1: 107710128-107710628, hg19; rs1938250: chr1: 120694097-120694597, hg19

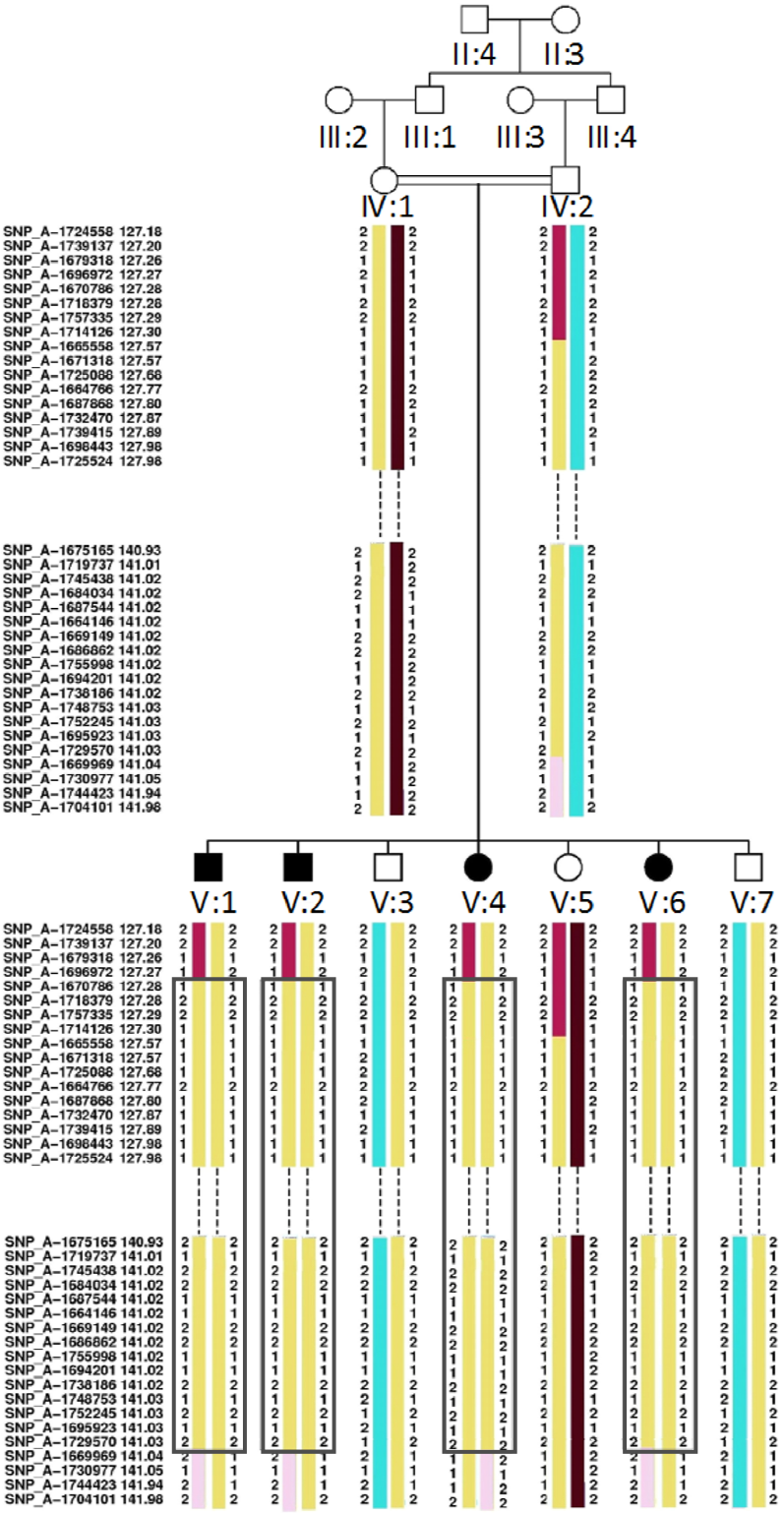

FIGURE 4.2: Haplotyping results for family M289. Grey frame: Homozygous haplotype between SNP markers rs10494061 and rs1938250 on chr.1p13.3–p11.2 in affected individuals.

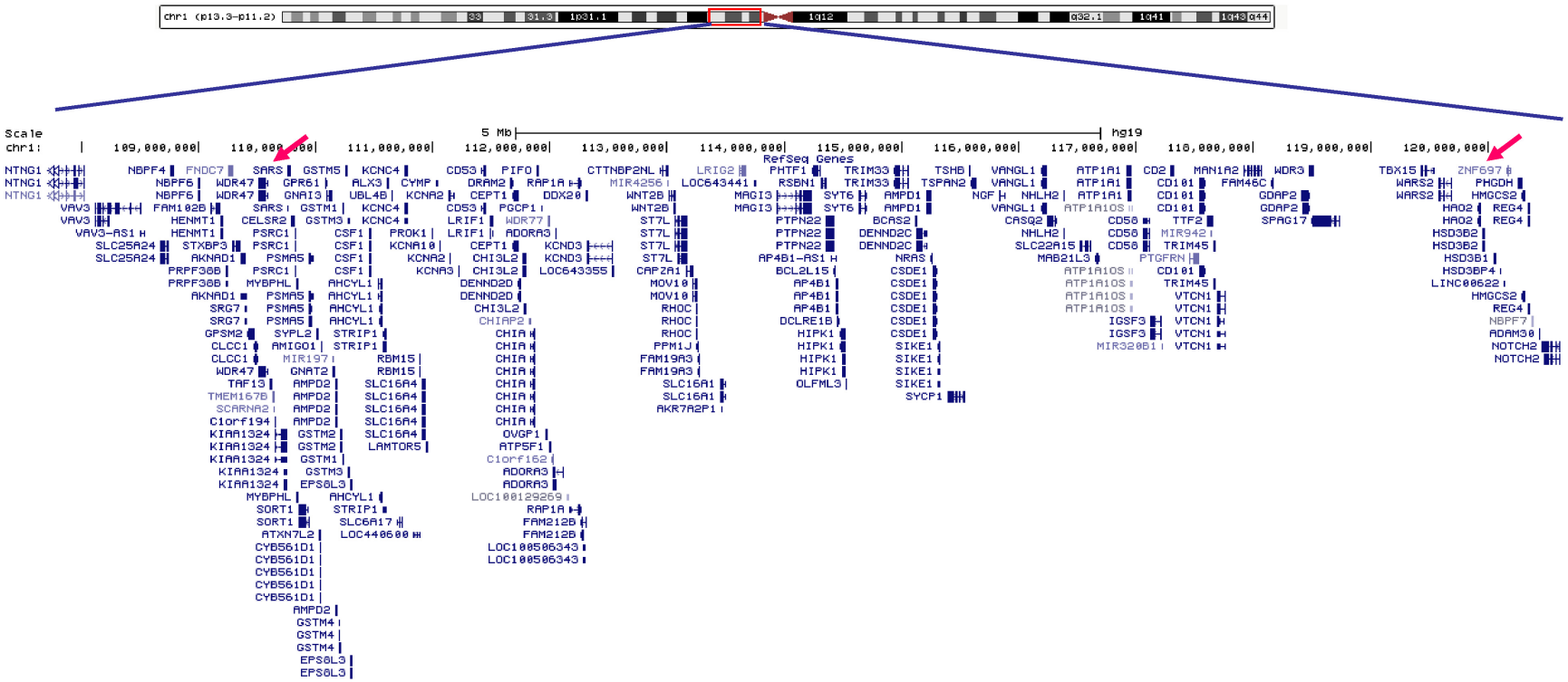

FIGURE 4.3: Genes in the linkage interval of family M289. RefSeq genes between SNPs rs10494061 and rs10494061 as depicted by the UCSC genome browser (hg19) (see section 3.3). Positions of *SARS* and *ZNF697* are marked by red arrows.

paragine[3] to aspartic acid at amino acid position 172 (p.D172N; NP_006504). In addition, a C > A point mutation (c.472C > A, chr1: 120166494; hg19) in exon three of *ZNF697* (zinc finger protein 697; NM_001080470) was found that results in a substitution of proline with threonine at amino acid position 158 (p.P158T; NP_001073939).

SARS c.514G > A was neither present in 420 population-matched healthy unrelated control individuals nor in 216 German controls. *ZNF697* c.472C > A was also not seen in the control panel (346 population-matched and 201 German controls).

Moreover, we used the publicly available sequencing data from the Exome Variant Server, 200 Danish individuals [Li *et al.*, 2010] and 185 genomes of healthy individuals made available by the 1000 Genome Project[4] [Abecasis *et al.*, 2010] to further extend the control cohort. Both missense mutations were not listed in these databases. The Exome Variant Server (October 2012) reports only one heterozygous nonsense mutation in *SARS* (c.319T > C [p.R107X]) at a frequency of 1 in 4406 in the African American population. Apart from this case, no other putatively deleterious mutations within the whole coding regions of *SARS* and *ZNF697* were reported.

4.2 Location of mutations in SARS and ZNF697 proteins

SARS encodes the cytosolic seryl-tRNA synthetase, a protein that contains three domains: an N-terminal tRNA binding arm, an aminoacylation domain and a C-terminal UNE-S domain containing the nuclear localization signal (NLS) [Xu *et al.*, 2012]. The amino acid substitution p.D172N found in family M289 is located in the aminoacylation domain, in close proximity to the active site.

The potential transcription regulator ZNF697 is a member of the cysteine$_2$ histidine$_2$ (C_2H_2) type zinc-finger protein family. ZNF697 contains 11 C_2H_2 zinc fingers. The first zinc finger domain has a near central position within the protein and lies separate from the other ten zinc finger domains, which are serially arranged close to the C-terminus. The P158T amino acid substitution is located upstream of the first krueppel-type zinc finger in a region without assigned function in both the Pfam and Interpro databases[5]. Schematic views of the gene and protein structures of both SARS and ZNF697, including the positions of the sequence alterations are depicted in Figure 4.4.

4.3 *SARS* and *ZNF697* are expressed during brain development

Expression of *SARS* and *ZNF697* was determined by RT-PCR in different human tissues including, most importantly, tissues of human brain regions relevant for learning and memory. Primers spanning the coding regions of *SARS* and *ZNF697* (primers see subsection 3.2.12) were used to amplify cDNA of both genes from RNA extracted from HeLa, HEK293-T, U373, SH-

[3] Three-letter and single-letter amino acid codes are listet in subsection 9.5.
[4] (see section 3.3)
[5] (see section 3.3)

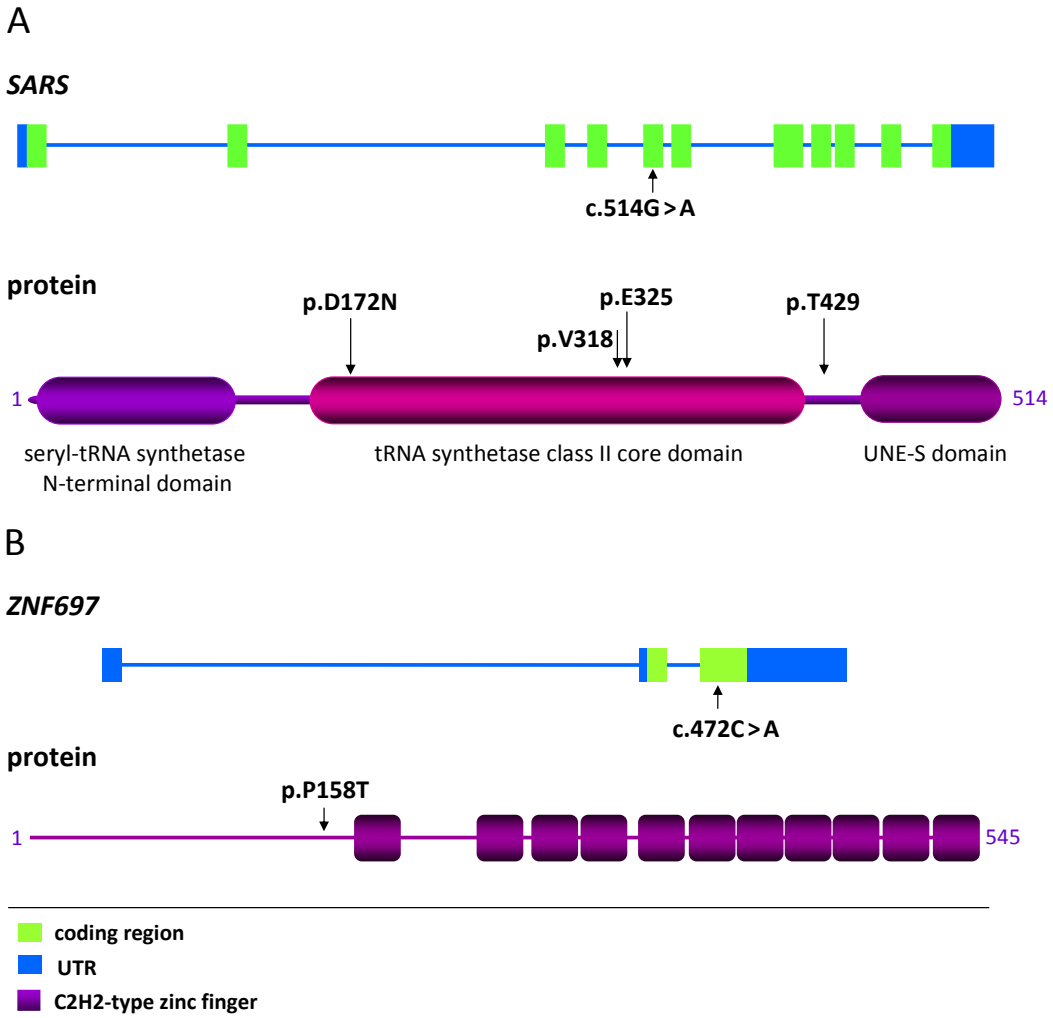

FIGURE 4.4: **A)** Schematic representations of the *SARS* gene and protein. The position of the missense mutation c.514G > A in exon five of *SARS* is indicated. Positions of the amino acid substitution (p.D172N), the ATP binding site (p.V318) and the serine binding sites (p.E325 and p.T429) are indicated above the domain structure (based on UniProtKB/Swiss-Prot entry P49591; see section 3.3). **B)** Schematic representations of *ZNF697* gene and protein. The position of missense mutation c.472C > A in the third exon of *ZNF697* is indicated by an arrow. The position of the amino acid substitution p.P158T is shown above of the schematic representation of the ZNF697 protein. Positions of C_2H_2-type zinc-fingers are based on UniProtKB/Swiss-Prot entry Q5TEC3.

SY5Y, lymphoblastoid and fibroblast cell lines as well as from template RNA derived from different adult and fetal human brain tissues (see subsections 3.2.7 and 3.2.8) (Figure 4.5). *SARS* and *ZNF697* were found to be well expressed in all brain tissues and all cell lines tested. Therefore, both genes may play a role during brain development and in normal brain functioning. Their expression in a broad range of cell lines renders these suitable model systems for studying the effects of the respective mutations *in vitro*.

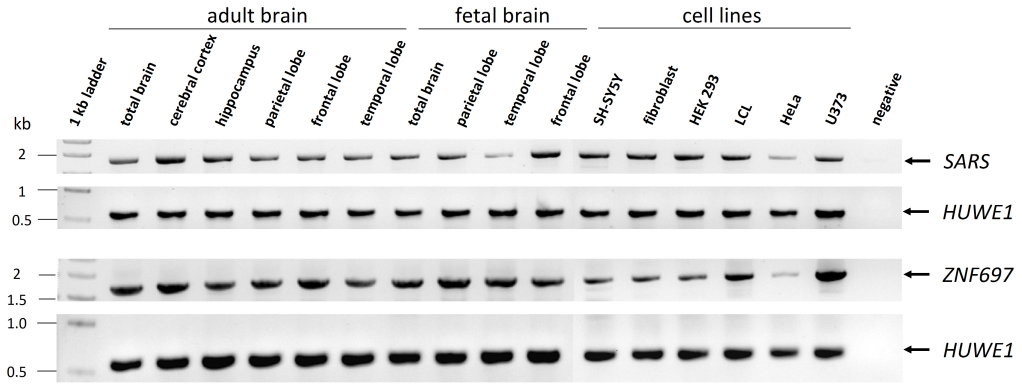

Figure 4.5: Expression of *SARS* and *ZNF697* in adult and fetal brain tissues and various human cell lines. Amplification of *HUWE1* was used to control success of cDNA synthesis.

4.4 Preparations for functional studies of SARS and ZNF697

Unfortunately, it was impossible to obtain any cell material from the affected individuals and healthy family members of family M289. Therefore, I set up several model systems to characterize the molecular functions of *SARS* and *ZNF697* and to determine their connection to the disease phenotype observed in the patients. For this purpose, I cloned the cDNAs of wild-type and mutant *SARS* and *ZNF697* in different expression vectors (see subsection 3.5.16). SARS was cloned in pGEX6P3 to obtain GST-SARS wild-type and mutant (p.D172N and p.T429A[6]) fusion proteins. The GST-tag allowed purification of the fusion proteins via Glutathione Sepharose chromatography (see section 3.6.1) for subsequent studies of SARS enzymatic activity. For immunohistochemistry experiments, the cDNAs of *SARS* and *ZNF697* were cloned into pEYFP-N1 and pmCherry-C1, to tag each protein both N-terminally and C-terminally with a fluorophore (Table 4.1).

4.5 Evolutionary conservation of SARS and ZNF697

SARS and *ZNF697* were both found to be expressed in fetal and adult brain. In order to evaluate the disease-causing potential of these genes, the degree of overall evolutionary conservation

[6] SARS p.T429A is enzymatically inactive, because the amino acid substitution affects a serine binding site. SARS p.T429A was used as a negative control during the pyrophosphate release assay and to detect potential differences in subcellular localization of ectopic wild-type and mutant SARS proteins in mammalian cell lines.

Table 4.1: *SARS* and *ZNF697* expression vectors used in functional studies

Plasmid	Properties
pEYFP-SARS-WT	pEYFP-N1 derivate containing 1542 bp *SARS* wild-type cDNA
pEYFP-SARS-D172N	pEYFP-N1 derivate containing 1542 bp *SARS* c.514G > A cDNA
pEYFP-SARS-T429A	pEYFP-N1 derivate containing 1542 bp *SARS* c1285A > G cDNA
pmCherry-SARS-WT	pmCherry-C1 derivate containing 1545 bp *SARS* wild-type cDNA
pmCherry-SARS-D172N	pmCherry-C1 derivate containing 1545 bp *SARS* c.514G > A cDNA
pmCherry-SARS-T429A	pmCherry-C1 derivate containing 1545 bp *SARS* c.1285A > G cDNA
pGEX6P3-SARS-WT	pGEX6P3 derivate containing 154 5 bp *SARS* wild-type cDNA
pGEX6P3-SARS-D172N	pGEX6P3 derivate containing 1545 bp *SARS* c.514G > A cDNA
pGEX6P3-SARS-T429A	pGEX6P3 derivate containing 1545 bp *SARS* c.1285A > G cDNA
pEYFP-ZNF697-WT	pEYFP-N1 derivate containing 1635 bp *ZNF697* wild-type cDNA
pEYFP-ZNF697-P158T	pEYFP-N1 derivate containing 1635 bp *ZNF697* c.472C > A cDNA
pmCherry-ZNF697-WT	pmCherry-C1 derivate containing 1638 bp *ZNF697* wild-type cDNA
pmCherry-ZNF697-P158T	pmCherry-C1 derivate containing 1638 bp *ZNF697* c.472C > A cDNA

of *SARS* and *ZNF* gene and protein sequences were assessed. Pairwise alignment scores were obtained from Homologene[7] (Table 4.2).

Table 4.2: Pairwise alignment scores of orthologs from other species and human *SARS* and *ZNF697*

	SARS	**Identity in [%]**		***ZNF697***	**Identity in [%]**	
	Gene symbol	**Protein**	**DNA**	**Gene symbol**	**Protein**	**DNA**
P. troglodytes	*SARS*	99.8	99.8	*ZNF697*	99.4	99.6
C. lupus	*SARS*	96.3	92.1	*ZNF697*	89.5	91.4
B. taurus	*SARS*	95.3	91.6	*ZNF697*	94.9	93.2
M. musculus	*Sars*	95.9	90.5	*Zfp697*	89.7	88.6
R. norvegicus	*Sars*	95.7	90.3	*Znf697*	89.5	88.0
G. gallus	*SARS*	85.8	78.4	–	–	–
D. rerio	*sars*	81.5	72.9	–	–	–
D. melanogaster	CG172559	70.0	66.7	–	–	–
C. elegans	*srs-2*	69.8	66.6	–	–	–
S. cerevisiae	*SES1*	52.1	53.0	–	–	–
N. crassa	NCU01443.1	52.2	55.6	–	–	–
A. thaliana	*AT5G27470*	52.7	57.1	–	–	–
P. falciparum	PF07_0073	49.8	50.9	–	–	–

For both, gene and protein alignment showed very high conservation between humans and chimpanzees. However, the degree of phylogenetic conservation of the SARS protein sequence between humans and rodents is much higher than the findings for ZNF697. Comparison of the protein sequences of SARS and ZNF697 between humans and mice, the latter being the premier mammalian model system for human diseases [Spencer, 2002], revealed that the proteins are 95.9 % and 89.7 % identical between species, respectively. Furthermore, human SARS is 81.5 %,

[7] (see section 3.3)

70.0 % and 69.8 % identical with cytoplasmic seryl-tRNA synthetases from the model organisms *D. rerio*, *D. melanogaster* and *C. elegans*, respectively. In summary, SARS is present throughout evolution and well conserved in eukaryota. ZNF697, however, is only conserved in eutherian mammals.

Next, the evolutionarily conservation of the affected SARS and ZNF697 protein residues was assessed by aligning protein sequences of SARS and ZNF697 orthologs from multiple species. The substituted p.D172 within SARS was conserved in all species except for yeast and bacteria. In contrast, p.P158 of ZNF697 was only conserved between rodents and humans (Figure 4.6).

Species	**SARS** **p.D172N** ↓	**ZNF697** **p.P158T** ↓
H. sapiens	YSHV**D**LVVMV	SRHRGDK**P**AHRRF
P. troglodytes	YSHV**D**LVVMV	SRHRGEK**P**AHRRF
C. lupus	YSHV**D**LVVMV	SRHRGDK**P**AHRRF
B. taurus	YSHV**D**LVVMV	GRHRGDK**P**AHRRF
M. musculus	YSHV**D**LVVMV	GRHRGDK**P**AYRRF
R. norvegicus	YSHV**D**LVVMV	SRHRSDK**A**AHRRF
D. rerio	YSHV**D**LVVMV	–
C. elegans	YSHV**D**LVVMV	–
A. thaliana	KNHV**D**LVELL	–
O. sativa	KNHV**D**LCKML	–
P. falciparum	YYHY**D**LLRKI	–
S. pombe	LSHH**E**VLTRL	–
M. grisea	LSHH**E**VLTRL	–
S. cerevisiae	LSHH**E**ILLRL	–
K. lactis	LSHH**E**ILLRL	–
E. gossypii	LSHH**E**VLLRL	–
N. crassa	LSHH**E**VLRKL	–

Figure 4.6: Multiple-species protein alignments of the affected amino acid positions within the SARS and ZNF697 proteins. The affected amino acids (bold and underlined) are shown together with flanking SARS or ZNF697 sequences and are marked by an arrow.

In line with this, the phyloP score, which measures the phylogenetic conservation of each nucleotide [Siepel *et al.*, 2006], was found to be 6.457 (i. e. very high conserved) for the relevant nucleotide (c.514G) of the *SARS* gene but only 0.395 (i. e. low conserved) for the *ZNF697* gene (c.472C), respectively.

4.6 *In silico* prediction of pathogenicity

The *SARS* and *ZNF697* genes are both well expressed in the brain and thus could both be implicated in disturbed intellectual functioning in the patients. To gain information about the individual disease causing potential of SARS p.D172N and ZNF697 p.P158T, I used five different algorithms[8] to predict the effect of each variant on protein function (Table 4.3).

[8] (see section 3.3)

Table 4.3: Predicted pathogenicity of SARS p.D172N and ZNF697 p.P158T

Programme	SARS p.D172N	ZNF697 p.P158T
PolyPhen2	probably damaging	benign
Mutation Taster	disease causing	polymorphism
PROVEAN	deleterious	neutral
SIFT	affect protein function	tolerated
PANTHER	deleterious	N.D.

N.D. not determined

All five programes unanimously predicted deleterious consequences for SARS function upon p.D172N substitution. In contrast, neither algorithm predicted a disease-causing potential for ZNF697 p.P158T[9]. These results further underscore that *SARS* is a plausible ARID candidate gene. Therefore, the subsequent experiments aimed at determining the potentially disease causing effects of p.D172N within SARS.

4.7 Aminoacyl-tRNA synthetases

SARS belongs to the class II family of aminoacyl-tRNA synthetases. Several genes encoding cytoplasmic and mitochondrial class I and class II aminoacyl-tRNA synthetases (ARSs) have important functions in neurons and have already been implicated in peripheral neuropathies and encephalopathies [Antonellis and Green, 2008; Konovalova and Tyynismaa, 2013], supporting the role of *SARS* as ID candidate gene.

ARSs are ancient enzymes and present in all organisms. At the beginning of the evolution of ARSs stands a core enzyme with the ability to activate amino acids. By insertions or fusions this enzyme gained binding sites for early tRNA-like oligonucleotides and thus obtained the ability for aminoacylation. Subsequently, the contemporary tRNA structure evolved and the ARSs aquired additional domains that added stability and enhanced specificity of each enzyme [Schimmel and Ribas De Pouplana, 2000]. There are two different classes of ARSs (class I and class II) which share neither any structural nor evolutionary relationship. Both classes are believed to have been present in the last common ancestor of archea, bacteria and eukaryota and each class contains ten enzymes [Schimmel, 2008]. Class I ARSs are mostly monomers and contain the two signature peptides histidine–isoleucine–glycine–histidine (HIGH) and lysine–methionine–serine–lysine–serine (KMSKS), important structures for aminoacylation at the active site. They constitute the so called Rossman fold that binds ATP and is responsible for the formation of aminoacyl adenylate [Schimmel, 2008]. Both motifs approach the end of the tRNA acceptor helix from the minor groove and catalyse the attachment of the cognate amino acid to the 2'-OH at the end of the tRNA chain [Cusack, 1997; Arnez and Moras, 1997; Cusack, 1999].

[9] *ZNF697* is an unlikely ARID candidate gene and might be a so far unreported SNP.

Most class II ARSs, including SARS, act as homodimers. The catalytic core of class II ARSs is composed of a seven-stranded β-structure with flanking α-helices [Schimmel, 2008]. Furthermore, these ARSs contain a set of three distinct motifs[10]: motif 1 is important for dimerization and several residues in motif 2 and 3 are involved in the binding of ATP and amino acid recognition [Ibba and Soell, 2000]. Class II ARSs approach the tRNA from the major groove and attach the amino acids to the 3' end of the tRNA [Wolf *et al.*, 1999].

4.7.1 The aminoacylation of tRNAs

Genetic information encoded by the mRNA is translated into the corresponding amino acid sequence, thus giving rise to protein molecules at the ribosome. The accuracy of this translation process is not only dependent on correct codon-anticodon matching between tRNA and mRNA at the ribosome but furthermore on the precision of the aminoacylation process, i. e. the charging of tRNAs [Ibba and Soell, 2000; Ibba *et al.*, 1997]. Aminoacylation is performed by the ARSs via a two-step reaction: First, an ARS binds an amino acid and ATP to form an aminoacyl adenylate (AA-AMP). At the same time, pyrophosphate ($P_2O_7^{4-}$, PP_i) is released (Equation 4.1). In the second step, the amino acid is transferred onto the 3' terminal adenosine of the tRNA that is to be charged and subsequently the charged tRNA and the AMP molecule are released from the enzyme (Equation 4.2; Figure 4.7) [Schimmel and Söll, 1979].

$$\text{ARS} + \text{AA} + \text{ATP} \longrightarrow \text{ARS(AA-AMP)} + \text{PP}_\text{i} \tag{4.1}$$

$$\text{ARS(AA-AMP)} + \text{tRNA} \longrightarrow \text{ARS} + \text{AA-tRNA} + \text{AMP} \tag{4.2}$$

As the genetic code is degenerate, ARSs need to identify and charge each of their several tRNA isoacceptors [Schimmel, 1991]. SARS needs to recognize six different $tRNA^{Ser}$ isoacceptors and the selenocysteine-tRNA ($tRNA^{Sec}$). This is achieved by interaction of the C-terminal coiled-coil domain of SARS with the long extra arm of all $tRNA^{Ser}$ isoacceptors and $tRNA^{Sec}$ [Biou *et al.*, 1994; Borel *et al.*, 1994].

4.8 The amino acid substitution p.D172N destabilizes SARS

Mutations of single amino acid mutations can significantly change the stability of a protein structure. p.D172 was predicted *in silico* to be a stabilization center element by the SCide software[11] using PDB structure 3VBB and chain F. To understand whether the stability of SARS protein is influcenced by the loss of the negative charge at residue 172, protein stability was assessed by MUpro, I-Mutant ($\Delta\Delta G$ = -2.06) and PoPMuSiC ($\Delta\Delta G$ = 0.62 kcal/mol). All three programmes[12] predicted a moderate decrease in protein stability upon mutation. Even more im-

[10] motif 1: +G(F/Y)XX(C/L/I)XXPhh, motif 2: +hhXhXXXFRXE and motif 3: hGhGhGhhERhhhh, where X stands for any amino acid, h for a hydrophobic and + for a positively charged residue [Carter, 1993].
[11] (see section 3.3)
[12] (see section 3.3)

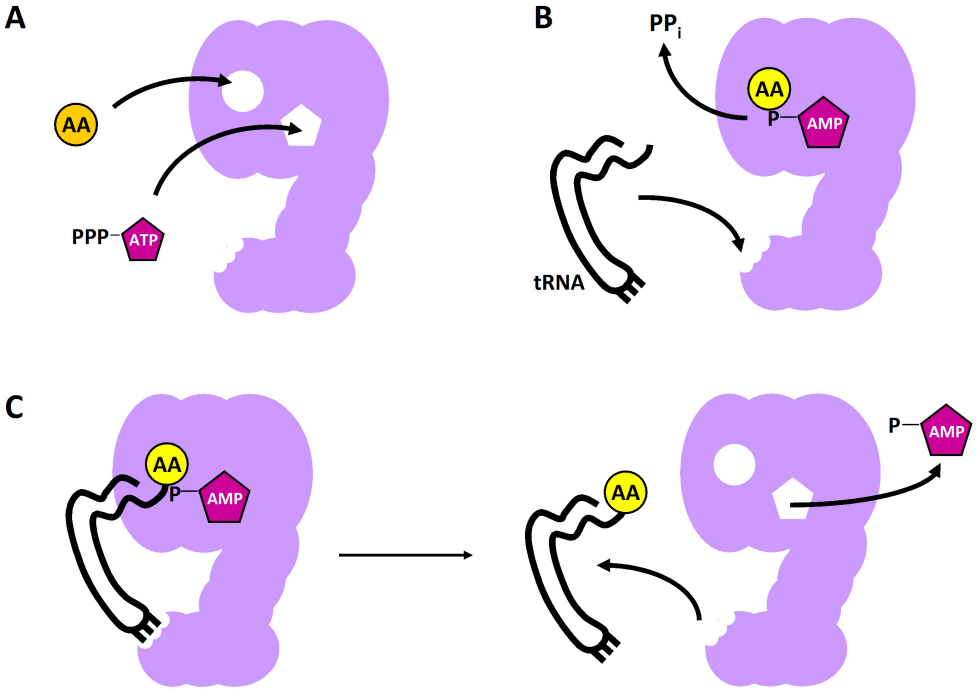

FIGURE 4.7: The process of aminoacylation. **A)** Amino Acid (AA) and ATP bind to the ARS. **B)** The amino acid is activated, aminoacyl-adenylate is formed and pyrophospahate (PP_i) is released. **C)** The tRNA binds and the amino acid is attached to the tRNA. Subsequently, the charged tRNA is released from the enzyme and is transferred to the ribosome. AMP is also released. Illustration according to [ANTONELLIS AND GREEN, 2008].

portant, *in silico* energy calculations with the Concoord/PBSA method [BENEDIX *ET AL.*, 2009] which calculates stability changes upon mutation confirmed a decrease in stability ($\Delta\Delta G$ = -1.06). Concoord/PBSA analysis was performed by Dr. H. Stehr.

4.9 Putative influence of p.D172N on the catalytic function and localization of SARS

To gain insights into the structural consequences of the p.D172N mutation, structure-based *in silico* modeling of the p.D172N amino acid substitution was performed by Dr. H. Stehr using crystal-structure 3VBB of human SARS [XU *ET AL.*, 2012]. p.D172 is located in the aminoacylation domain, in close proximity to p.H170 (Figure 4.8). XU *ET AL.* have shown that residues p.H170 and p.F316 form hydrophobic interactions to stabilize the β10-β11 hairpin. This hairpin belongs to the seven-stranded antiparallel β-sheet (β1-β9-β10-β11-β13-β8-β7) which is the central core of the aminoacylation domain [XU *ET AL.*, 2012].

Aspartic acid and asparagine are geometrically similar and p.D172 faces away from the active site, therefore p.D172N has probably no structure-based influence on the catalytic core domain. However, p.D172 could affect aminoacylation activity in an indirect manner as it provides a

negative charge. This negative charge is lost in SARS p.D172N. The results of the above analysis for wild-type and mutant SARS residues are shown in Figure 4.8.

Furthermore, modeling revealed that the mutation site p.D172 is located directly adjacent to the N-terminus of the UNE-S domain, which contains the NLS. Negative charges within the UNE-S domain affect the stability of the positively charged NLS [Xu *et al.*, 2012]. Therefore, the substitution of the negatively charged p.D172 to the the positively charged p.N172 could influence the subcellular localization of the SARS protein.

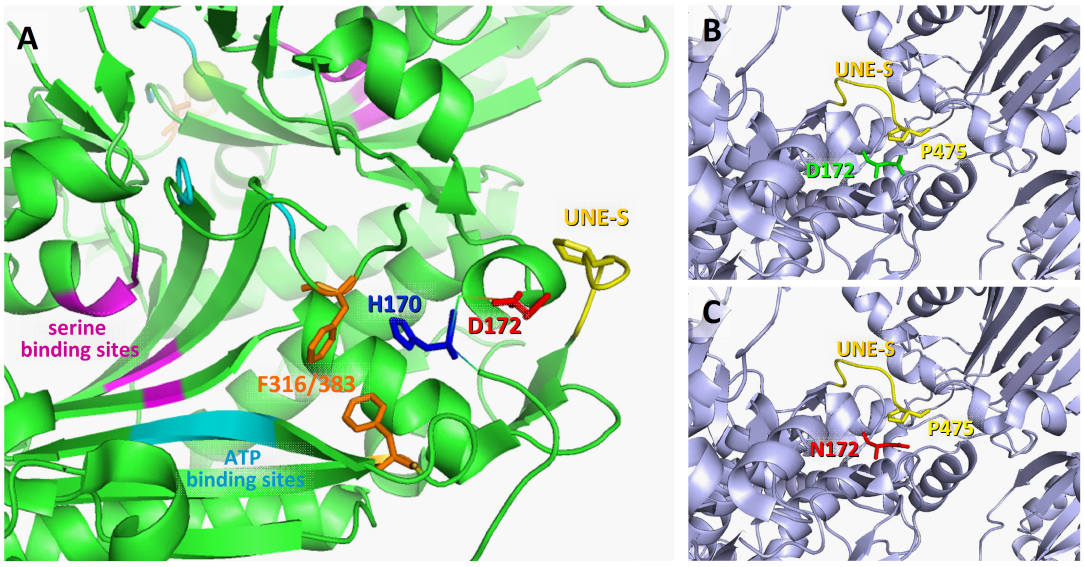

FIGURE 4.8: **A)** Active site of the SARS wild-type protein. Residues p.H170 (dark blue) and p.F316/383 (orange) stabilize a β-sheet of the aminoacylation domain. p.D172 (red) faces away from the aminoacylation core, but is located directly adjacent to the N-terminus of the UNE-S domain (yellow). Light blue: ATP binding sites; pink: serine binding sites. **B)** Orientation of wild-type p.D172 (green) and **C)** orientation of mutant p.N172 (red) protein residues. The nuclear localization signal (NLS) is located seven amino acids downtream of p.P475.

4.10 Subcellular localization of SARS p.D172N

To test the hypothesis that p.D172N might influence the localization pattern of SARS, expression of wild-type and mutant SARS was studied in three different cell lines. SH-SY5Y cells were used to set up a neuronal model system and HeLa and HEK293-T cells were chosen as basic cell biologcal models. First, subcellular localization of endogenous SARS was analysed using a monoclonal SARS antibody (see subsection 3.2.11).

In all three cell lines endogenous SARS was detected in the cytoplasm (Figures 4.9 and 9.1). Next, expression vector systems using N- as well as C-terminally tagged SARS wild-type and mutant (p.D172N and T429A) constructs were set up (Figures 4.10, 4.11 and 4.12). SARS p.T429A is enzymatically inactive and was used as an additional control for detection of potential differences in subcellular localization of ectopic wild-type and mutant SARS proteins.

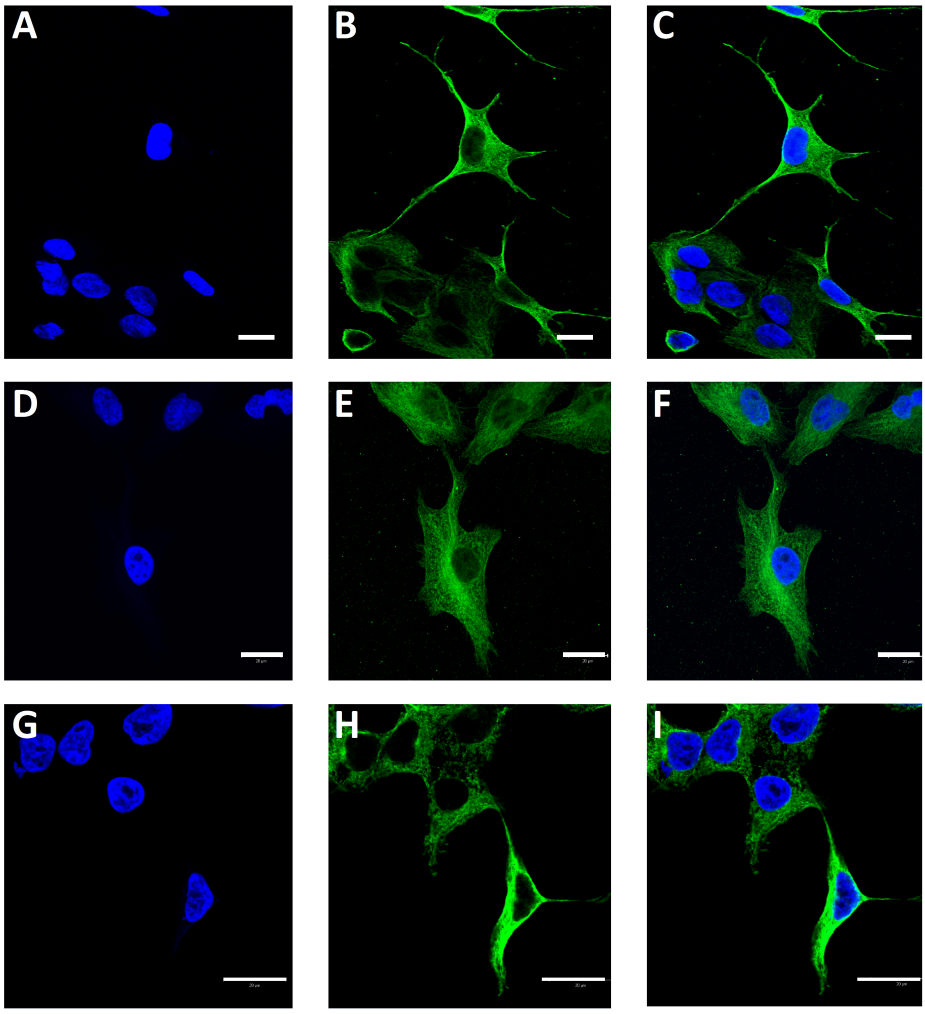

FIGURE 4.9: Confocal immunofluorescence microscopy showing the cytoplasmic localization of endogenous SARS in various cell lines, **(A–C)** SH-SY5Y, **(D–F)** HEK293-T, and **(G–I)** HeLa. The green signal corresponds to SARS staining, whereas the blue corresponds to nuclear DAPI staining. Scale bars = 20 µm.

Confocal micoscopy reavealed cytoplasmic localization of endogenous and ectopic SARS proteins in SH-SY5Y, Hela and HEK293-T cells[13]. These findings are in agreement with the confocal micoscopy studies by XU *ET AL.* who report that indeed the bulk of SARS localizes to the cytoplasm. Furthermore, they also observed minor fractions of SARS to be present in the nucleus where SARS shows a pronounced spotted pattern. Notably, their micoscopic evidence of nuclear SARS was restricted to expression in HUVEC (Human Umbilical Vein Endothelial Cells) [XU *ET AL.*, 2012].

[13] Results of confocal immunofluorescence microscopy showing the cytoplasmic localization of SARS p.T429A are shown in section 9.2

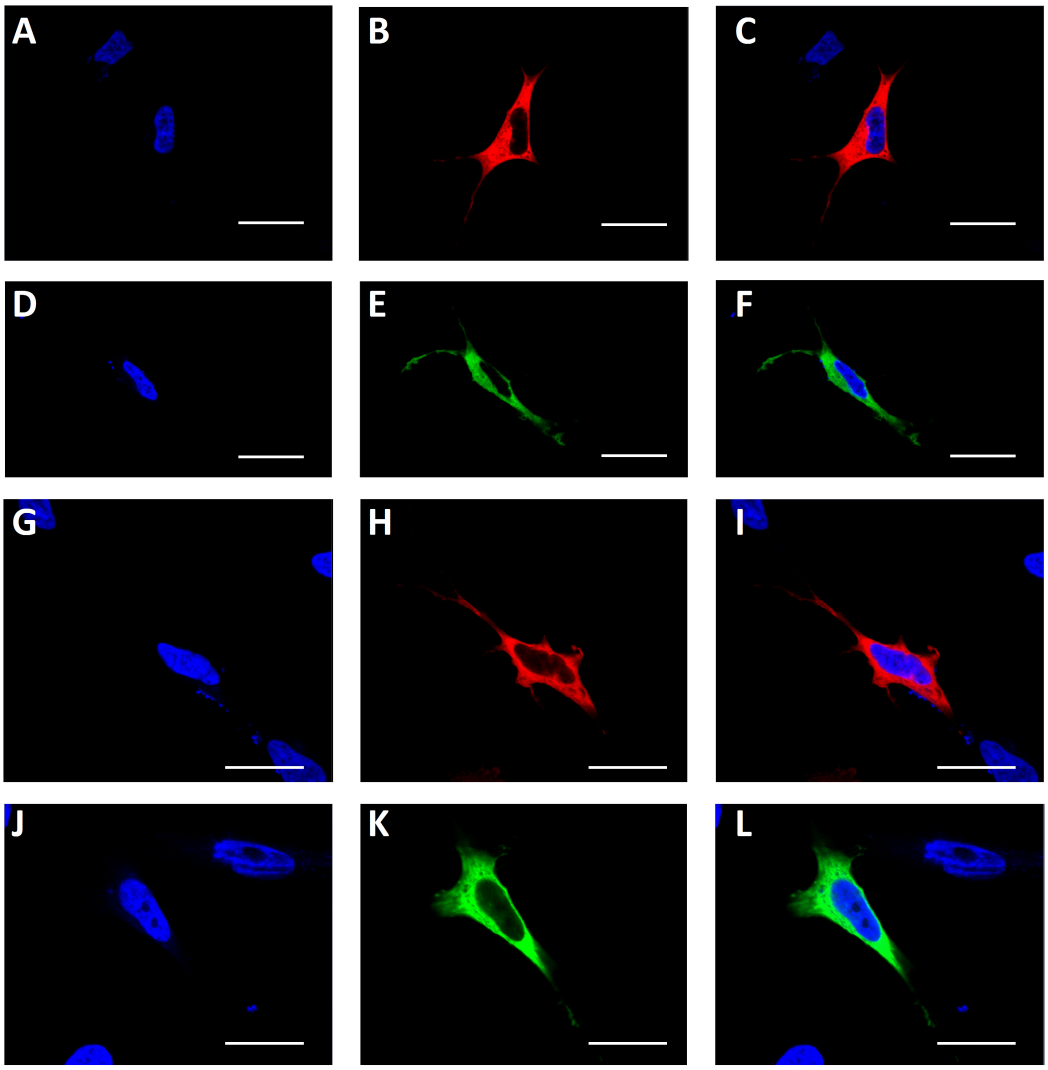

FIGURE 4.10: Confocal immunofluorescence microscopy showing the cytoplasmic localization of **(A–F)** wild-type and **(G–L)** p.D172N SARS-EYFP and SARS-mCherry in SH-SY5Y cells. Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

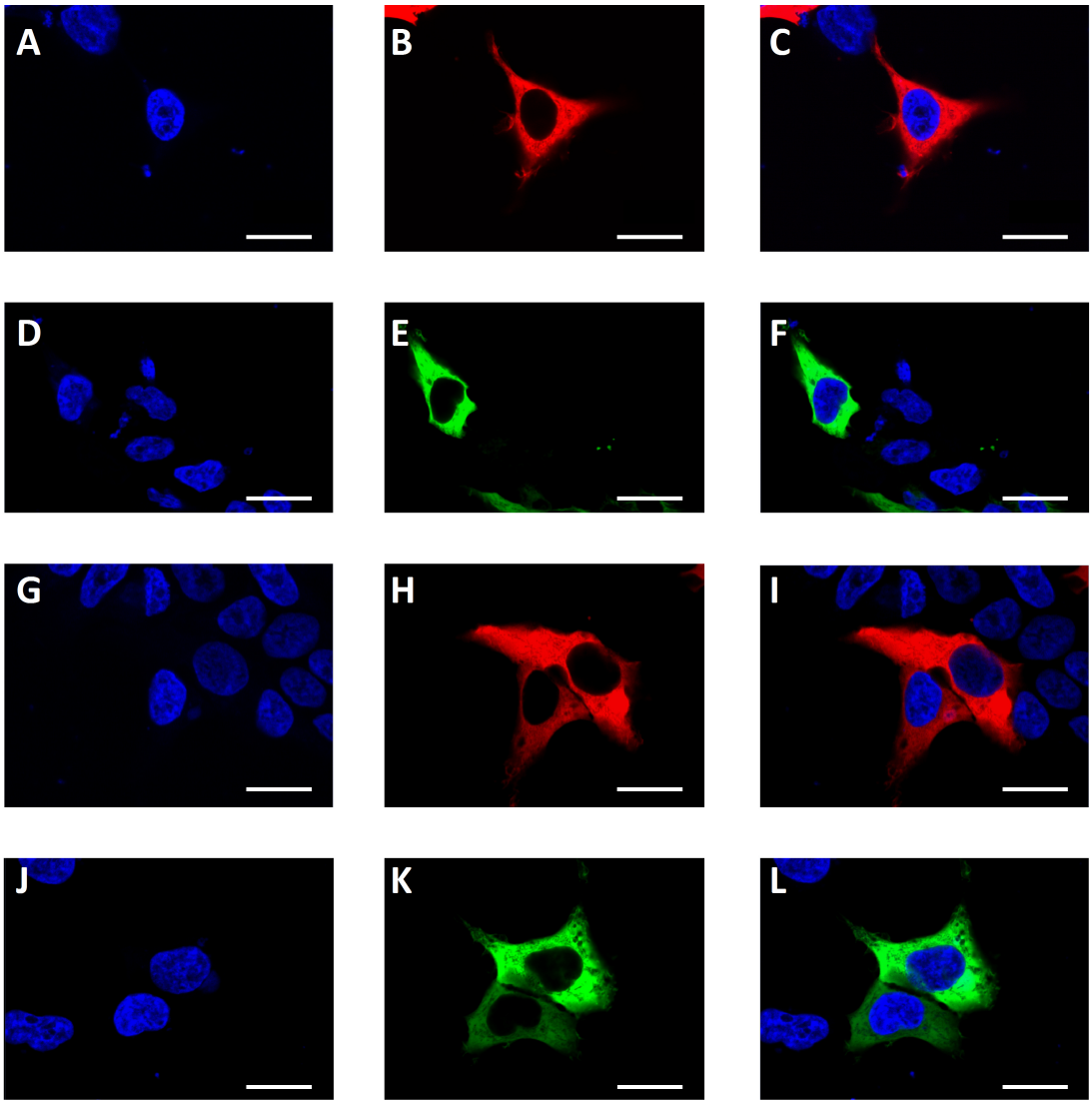

FIGURE 4.11: Confocal immunofluorescence microscopy showing the cytoplasmic localization of **(A–F)** wild-type and **(G–L)** p.D172N SARS-EYFP and SARS-mCherry in HEK293-T cells. Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

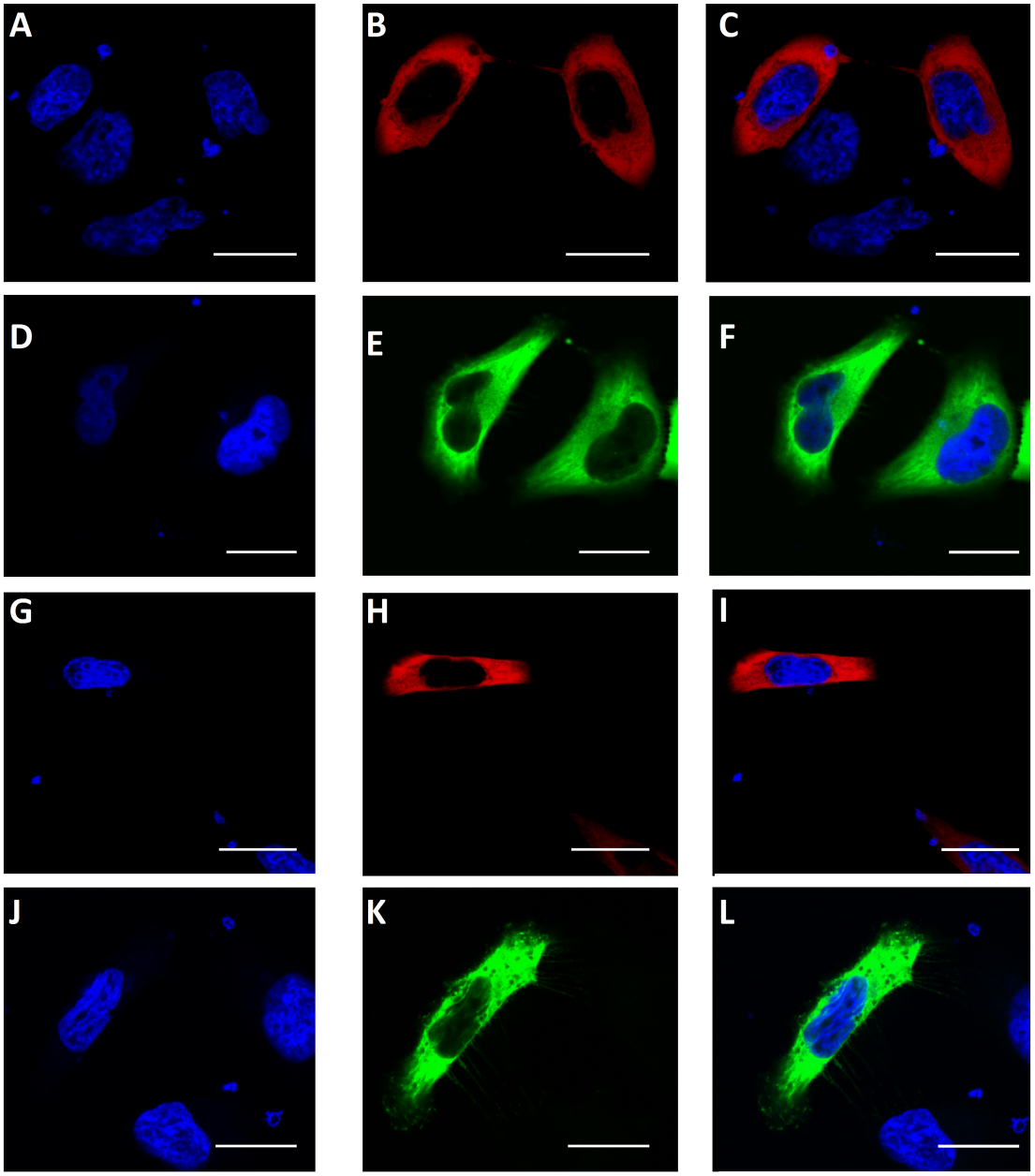

FIGURE 4.12: Confocal immunofluorescence microscopy showing the cytoplasmic localization of **(A–F)** wild-type and **(G–L)** p.D172N SARS-EYFP and SARS-mCherry in HeLa cells. Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

4.11 Expression analysis of SARS p.D172N in mammalian cells

HEK293-T cells were transfected with *EYFP-SARS* wild-type and p.D172N expression vectors for 24 hours and transfection efficiency was determined by counting transfected cells by fluorescence microscopy. Transfections rates were 40.8 % ± 3.2 % for SARS wild-type and 39.8 % ± 1.1 % for SARS p.D172N, respectively. Data were derived from three independent transfections.

After transfection, cells were harvested and fractionated (see section 3.6.9). Expression of ectopic wild-type and mutant SARS proteins in whole cell lysate and cytosolic as well as nuclear extracts was analysed by Western Blotting using monoclonal anti-GFP antibody. The same membranes were then hybridized with anti-tubulin and anti-lamin antibody. Tubulin and lamin were used as control for cell fractionation and for normalization of ectopic SARS protein expression (Figures 4.13 and 4.14).

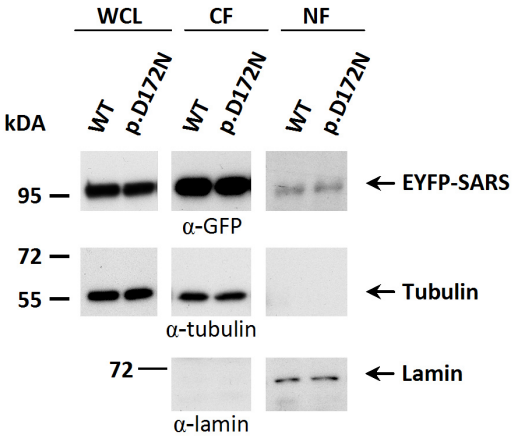

Figure 4.13: Expression of ectopic SARS proteins. Whole cell lysate (WCL), cytoplasmic (CF) and nuclear (NF) fractions from HEK293-T cells transfected either with wild-type or mutant EYFP-SARS were run on SDS-PAGE. The gel was blotted and probed with α-GFP antibody. The blot was subsequently probed with α-tubulin and α-lamin as loading controls for cytosolic and nuclear fractions, respectively.

The concentration of ectopic SARS p.D172N protein was reduced to 67 % as compared to ectopic SARS wild-type protein in whole cell extracts ($P < 0.005$) as well as in the cytosolic fraction ($P < 0.05$) (Figure 4.14). In the nuclear fraction no statistically significant differences in protein concentration could be observed. Normalization of EYFP-SARS bands to lamin bands revealed SARS/lamin ratios of 2.4 % ± 0.5% for EYFP-SARS wild-type (n=6) and 1.7 % ± 1.0% for EYFP-SARS p.D172N (n=6), respectively. However, quantification of protein amounts in the nuclear fraction was complicated due to low protein expression.

4.12 Impaired serine-activation of the SARS p.D172N enzyme

In silico modeling of SARS revealed that p.D172 maps close to the active site of the protein (Figure 4.8). Although it is facing away from the active site, p.D172 might influence SARS aminoacylation activity, because of its negative charge, which is lost upon mutation (p.D172N). To un-

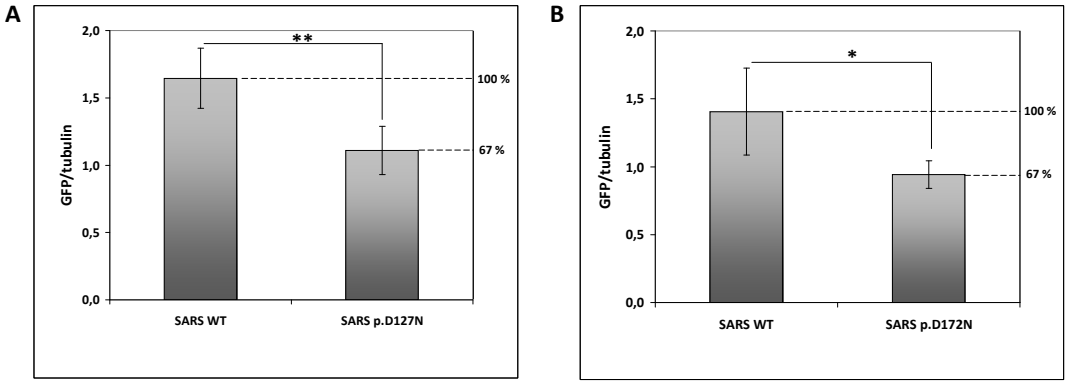

Figure 4.14: Quantification of wild-type and mutant EYFP-SARS expression in HEK293-T cells. Quantification of bands from Western blots was performed using ImageQuant software (Molecular Dynamics). Tubulin was used to normalize SARS protein expression. **A)** Histogram showing a statistically highly significant reduction of SARS p.D172N expression compared to expression of SARS wild-type in total cell lysate from HEK293-T cells (n=8; *t* test (two-tailed, homoscedastic): **, $P < 0.05$). SARS p.D172N expression is reduced to 67 % of SARS wildtype expression. **B)** Histogram showing significant reduction of mutant SARS in the cytosolic fraction as compared to wildtye SARS (n=6; *t* test (two-tailed, homoscedastic): **, $P < 0.05$). SARS p.D172N expression is reduced to 67 % of SARS wildtype expression.

derstand whether p.D172N SARS is still able to perform the first step of aminoacylation (see section 4.7) a pyrophosphate release assay was performed with wild-type SARS, p.D172N SARS and p.T429A SARS. The latter is unable to bind serine as p.T429A affects a serine binding site (Figure 4.4) and was used as a negative control during the aminoacylation test.

4.12.1 Overexpression and purification of wild-type and mutant SARS

SARS wild-type and mutant (p.D172N and p.T429A) cDNAs were cloned into pGEX6P3. Expression of GST-SARS fusion proteins in *E. coli* BL21 was induced with IPTG for four hours. Subsequently, GST-SARS fusion proteins were purified using glutathione sepharose columns. Glutathione was removed from the eluted proteins by buffer exchange. Quality of protein overexpression, purification and buffer exchange were checked by SDS-PAGE and subsequent Western Blotting (Figure 4.15).

All four proteins were expressed in *E. coli*. Fusion proteins GST-SARS wild-type, GST-SARS p.T429A and the unfused GST were stable thoughout purification and buffer exchange. In contrast, GST-SARS p.D172N was very unstable during these procedures and major portions of the protein were degraded so that only a low amount of GST-SARS p.D172N could be obtained. Degradation of the p.D172N SARS protein has been observed during all purification experiments performed during this study. The purification conditions have been optimized so that a sufficient amount of SARS p.D172N for experimental use could be obtained. The SDS polyacrylamid gel was then stained with silver to quantify various proteins using a serial dilution of BSA as a standard (Figure 4.16).

Quantification of SARS protein bands was performed using the ImageQuant software (Molecular Dynamics). GST-SARS wild-type and p.T429A concentrations were 0.35 ng/µl and 0.30 ng/µl

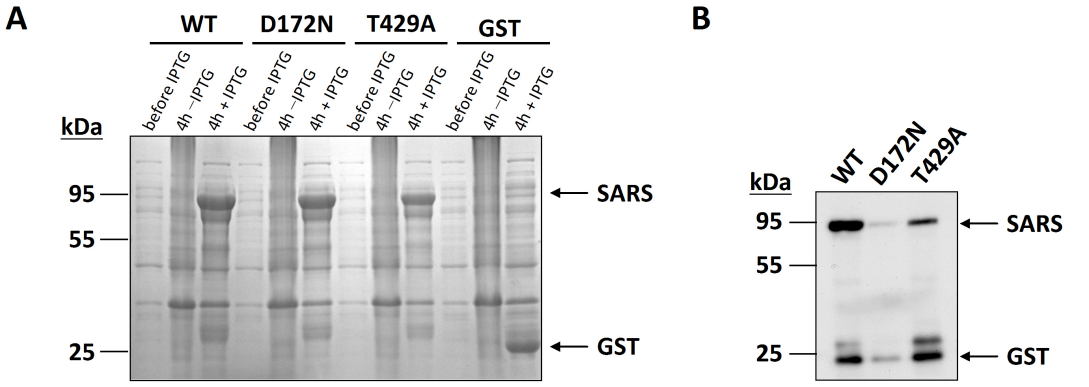

Figure 4.15: **A)** SDS-PAGE of bacterial lysates stained with Imperial™Protein Stain (Thermo Scientific). Lysates of bacteria transfected with plasmids containing either GST-SARS wild-type, p.D172N, p.T429A or unfused GST before induction with IPTG and after incubation for four hours with or without 1 mM IPTG were loaded. After IPTG induction SARS-GST proteins show a clear band at 90 kDA. **B)** Western Blot showing purified GST-SARS fusion proteins. Equal volumes of proteins were loaded. The molecular weight is given in kDA. Proteins were detected by α-GST. The position of GST-SARS fusion protein bands (84 kDa) and GST protein band (26 kDa) are indicated.

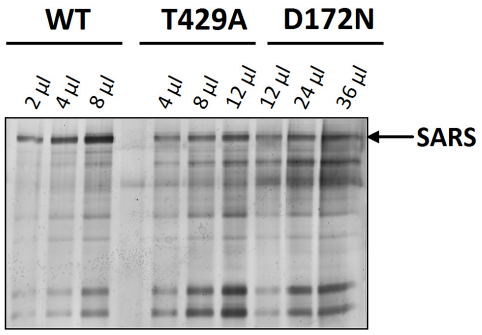

Figure 4.16: SDS-PAGE of GST-SARS fusion proteins after buffer exchange. The SDS-gel was stained using ProteoSilver™ Plus Silver Stain Kit (Sigma). Protein volumes of GST SARS wild-type, p.D172N and p.T429A loaded onto the gel are given. The position of GST-SARS fusion proteins (84 kDa) is indicated by an arrow.

respectively. The concentration of GST-SARS p.D172N was seven times lower (0.5 ng/µl) as compared to GST-SARS wild-type protein. Concentration of GST (0.3 µg/µl) was determined with a spectrophotometer.

4.12.2 SARS p.D172N displays impaired serine activation activity *in vitro*

The enzymatic function of SARS is to charge its cognate tRNAs with serine. During the first step of this reaction, serine-adenylate is formed and pyrophosphate is released (Equation 4.3). Subsequently, pyrophosphate hydrolyses into inorganic phosphate (Equation 4.4).

$$\text{SARS} + \text{serine} + \text{ATP} \longrightarrow \text{SARS(serine-AMP)} + P_2O_7^{4-} \tag{4.3}$$

$$P_2O_7^{4-} + H_2O \longrightarrow 2\,HPO_4^{2-} \tag{4.4}$$

I studied the ability of the SARS wild-type and mutant (p.D172N and p.T429A) proteins to form seryl-adenylate in a pyrophosphate release test. Enzyme activity was assayed at 37 °C and free phosphate was detected using BIOMOL GREEN Reagent™ as described (see section 3.8.1). The purified GST protein was used as an additional negative control. For each protein, activity measurements were performed in four and six probes at three different time points (0, 60 and 90 minutes) (Figure 4.17).

The pyrophosphate release of the wild-type SARS was 100 % at 60 minutes and after 90 minutes of incubation the enzyme still released 67 % pyrosphosphate as a by-product of serine activation. This reduced rate of pyrophosphate production can be explained by entering of AMP into the back reaction of amino acid activation, product inhibition by the generated seryl-adenylate, or inhibition of SARS by pyrophosphate (discussed in section 6.1.5).

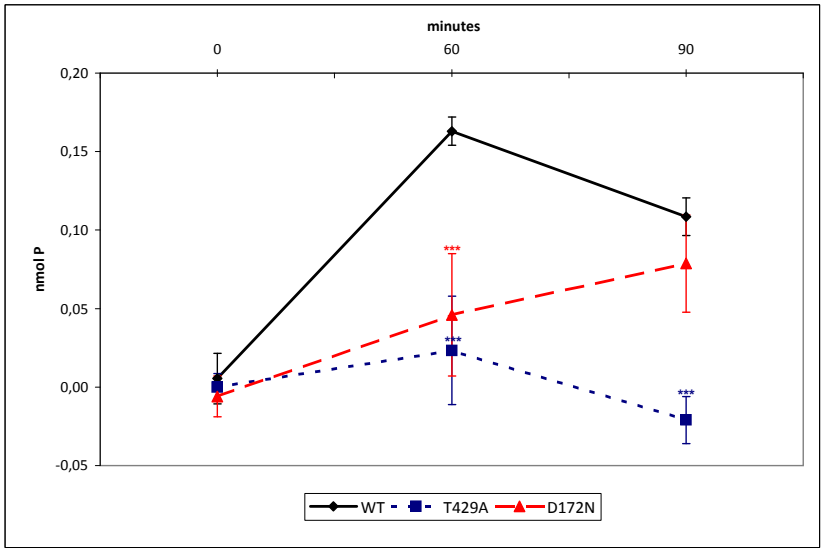

FIGURE 4.17: Results of pyrophosphate release assay. Equal amounts of recombinant SARS wild-type (WT) and mutant (p.D172N and p.T429A) proteins were used to assay the first step of aminoacylation. The pyrophosphate released (nmol P) during serine activation was measured in a colorimetric reaction using BIOMOL GREEN Reagent™. At 60 minutes p.D172N SARS (n = 4; red line) shows highly statistically significant reduction of pyrophosphate release as compared to wildtype SARS. The catalytically inactive p.T429A SARS (blue line) released no pyrophosphate (n=6). Data were normalized to individual background activity. *** $P < 0.001$ (Student's t test; two-tailed, homoscedastic).

Remarkably, a highly significant loss of serine-activation was observed for the SARS p.D172N mutant. At 60 minutes GST-SARS p.D172N released only 28 % pyrophsophate as compared to the GST-SARS wild-type enzyme and even after 90 minutes of incubation GST-SARS p.D172N still produced 51 % less pyrophsophate as the wild-type enzyme (release at 60 minutes). The GST-SARS p.T429A mutant is unable to perform the aminoacylation reaction because a serine binding site is destroyed by this amino acid substitution. Consistent with previous findings (e. g. XU *ET AL.* [2012]) the GST-SARS p.T429A was inactive during the assay. As expected, the GST protein did not show any pyrophsophate release (data not shown).

5 Results – Family M8600485

5.1 Linkage analysis and mutation screening in family 8600485

The affected individuals of family 8600485 all presented with a pronounced phenotype (see section 3.1.2). Parametric linkage analysis in family 8600485 revealed a single interval of homozygosity on chromosome six with a LOD score of 3.1 (Figures 5.1 A and 5.2)[1]. The linkage interval had a size of 9.7 Mb was flanked by the heterozygous SNPs rs9379512 (chr6:23391399-23391899, hg19) and rs10484569 (chr6:33058702-33059202, hg19) (Figure 5.2).

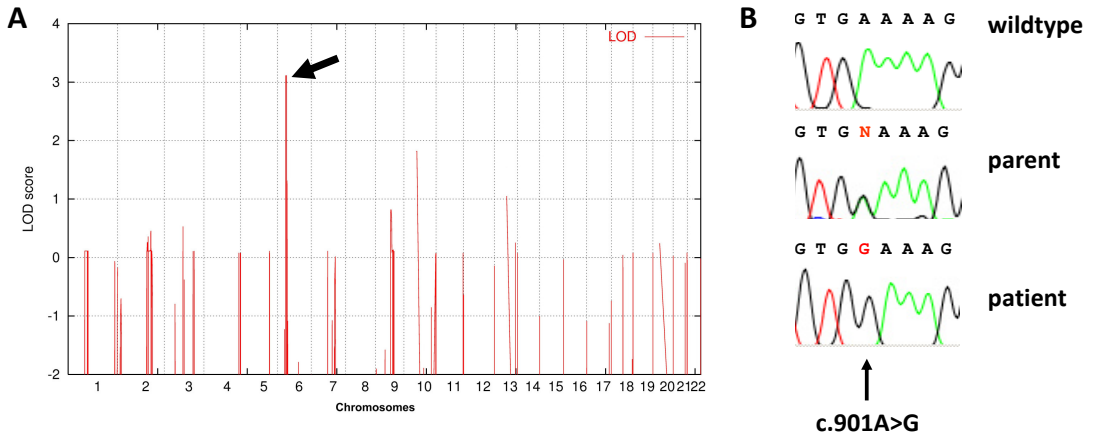

FIGURE 5.1: **A)** Linkage profile resulting from whole genome parametric linkage analysis, showing a single linkage interval (arrow) with significant LOD score 3.1 on chromosome 6p22-p21.32. **B)** Sequence chromatograms of an affected individual, a parent and a healthy control (wild-type), showing the homozygous substitution c.901A > G in *ALDH5A1* in the affected individual.

The gene *ALDH5A1* (aldehyde dehydrogenase 5a1, NM_001080, MIM #610045) was one of the 284 RefSeq genes (hg19) located within this interval (Figure 5.3). *ALDH5A1* is a known disease gene and has been associated with succinic semialdehyde dehydrogenase (SSADH) deficiency (MIM #271980), an autosomal recessive inherited neurodevelopmental disorder. As the patients' phenotype showed common features of SSADH-deficiency, in this study, all coding exons and exon-intron boundaries of *ALDH5A1* were amplified by PCR and subsequently analyzed by Sanger sequencing. This led to the identification of an A > G substitution in exon six

[1]Linkage analysis was carried out by Dr. M. Garshasbi. Genotyping (SNP analysis) of all affected family members, their parents, healthy sibling and the affected cousin was performed using the Human 610-Quad BeadChip (Illumina) following the protocol of the manufacturer. Details of data quality controls and linkage analysis have been published elsewhere [GARSHASBI *ET AL.*, 2006].

of *ALDH5A1* (c.901A > G, chr1: 24520659, hg19) which leads to a substitution of leucine[2] with glutamic acid at amino acid position 301 in the SSADH protein (p.K301E; NP_001071). This missense mutation co-segregated with the disease in the core family and was also present in the affected cousin. c.901A > G was not found in 94 population matched healthy unrelated control individuals as well as 124 German controls. Moreover, c.901A > G was not reported in any of the genome databases[3] used to extend the control cohort.

5.2 The p.K301E mutation resides near an active site of SSADH

ALDH5A1 encodes SSADH, a NAD^+ dependent mitochrondrial matrix enzyme. SSADH is involved in the final degradation step of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA), which results in the conversion of succinic semialdehyde (SSA) into succinic acid (succinate) in the mitochondrial matrix. The substrate binding sites are located at protein residues p.R213, p.R334 and p.S498. The active sites of SSADH are located at protein residues p.E306 and p.C340. The missense mutation resulting in p.K301E is thus located in close proximity to the first of these active sites. Schematic views of the gene and protein structures of both *ALDH5A1* and SSADH, including the position of the mutation are depicted in Figure 5.4.

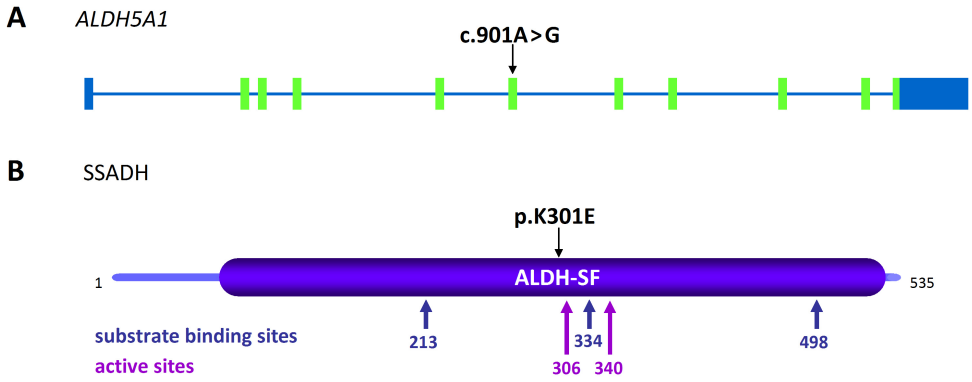

FIGURE 5.4: **A)** Schematic representations of *ALDH5A1*. The position of missense mutation c.901A > G in exon six of *ALDH5A1* is indicated. **B)** Schematic representation of SSADH protein. The relative positions of the substituted p.K301 is labeled on top of the domain structure and the substrate binding sites (blue) and the active sites (violet) are indicated below the domain structure. ALDH-SF: aldehyde dehydrogenase superfamily domain. Positions of the active sites and the substrate binding sites are based on `http://www.uniprot.org/uniprot/P51649`.

5.3 p.K301E within SSADH has a high disease causing protential

Next, evolutionary conservation of the *ALDH5A1* gene, its gene product, the SSADH protein, and both the affected nucleotide and substituted amino acid were investigated. *In silico* analysis

[2] Three-letter and single-letter amino acid codes are listet in subsection 9.5.
[3] 1. Exome Variant Server; 2. 200 Danish individuals [LI *ET AL.*, 2010]; 3. 1000 Genome Project; references see above

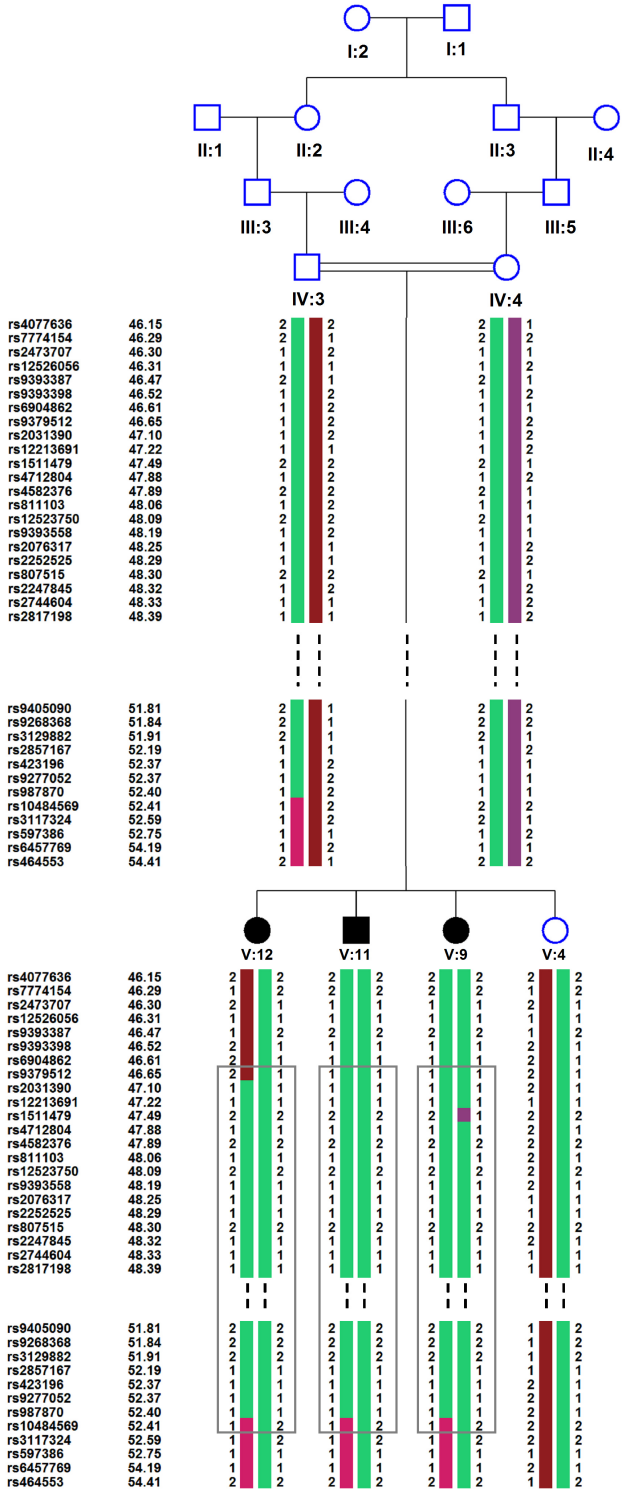

FIGURE 5.2: Haplotyping results for family 8600485. Grey frame: Homozygous haplotype between SNP markers rs9379512 and rs10484569 on chr.6p22-p21.32 in affected individuals.

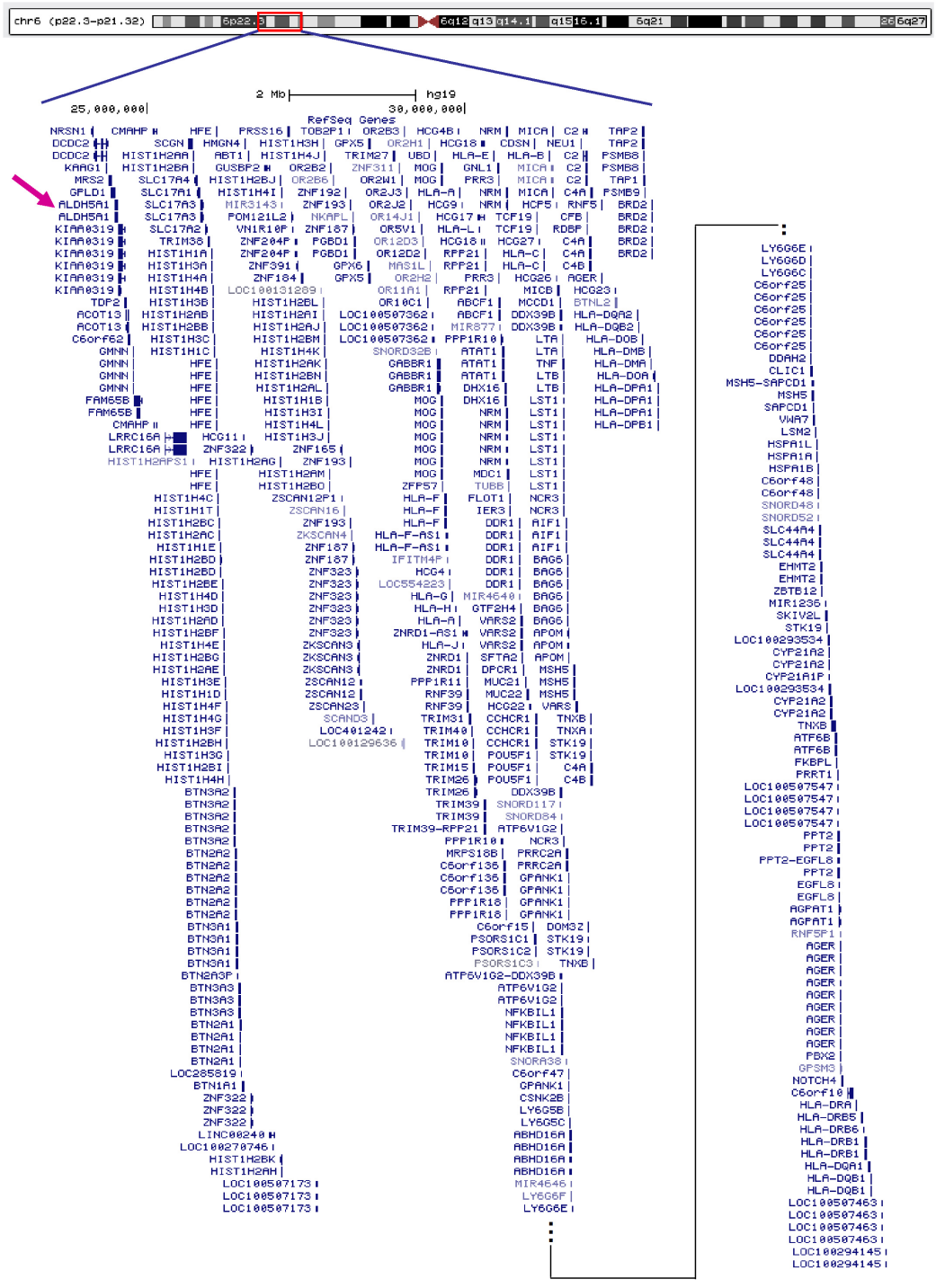

FIGURE 5.3: Genes in the linkage interval of family 8600485. RefSeq genes between SNPs rs9379512 and rs10484569 as depicted by UCSC genome browser (see section 3.3) (hg19). *ALDH5A1* is marked by a red arrow.

with Homologene[4] revealed that SSADH is highly evolutionarily conserved and is present in all eukaryota (Table 5.1). *ALDH5A1* c.901A has a phyloP score of 5.031 (i. e. conserved) and p.K301 is highly conserved throughout the animal kingdom (Figure 5.5).

Table 5.1: Pairwise alignment scores of orthologs from other species and human *ALDH5A1*

	ALDH5A1	**Identity in [%]**	
	Gene symbol	**Protein**	**DNA**
P. troglodytes	*ALDH5A1*	99.6	99.6
C. lupus	*ALDH5A1*	91.4	87.7
B. taurus	*ALDH5A1*	85.9	86.1
M. musculus	*Aldh5a1*	86.2	82.0
R. norvegicus	*Aldh5a1*	86.0	82.0
G. gallus	*ALDH5A1*	78.4	72.6
D. rerio	*aldh5a1*	70.7	67.6
D. melanogaster	*Ssadh*	53.3	55.1
A. gambiae	AgaP_AGAP003165	54.2	56.6
C. elegans	*alh-7*	50.1	53.7
S. cerevisiae	*UGA2*	48.7	53.6
K. lactis	KLLA0E17491g	50.6	54.3
N. crassa	NCU00936	57.5	57.9
A. thaliana	*ALDH5F1*	59.5	58.6

In further *in silico* analyses to determine the disease causing potential of *ALDH5A1* c.901A > G, all five of the programmes used unanimously predicted deleterious consequences for SSADH function (PolyPhen2: "probably damaging"; Mutation Taster: "disease causing"; SIFT: "mutation affects protein function"; PANTHER: "deleterious"; PROVEAN: "deleterious").

Species	**p.K301E** ↓
H. sapiens	NSV**K**RVSMEL
P. troglodytes	NSV**K**RVSMEL
M. mulatta	NSV**K**RVSMEL
M. musculus	NSV**K**RVSMEL
R. norvegicus	NSV**K**RVSMEL
B. taurus	NSV**K**RVSMEL
C. lupus	GSV**K**RVSMEL
G. gallus	GTV**K**RVSMEL
D. rerio	GTV**K**RVSMEL
C. eleangs	STV**K**RVCLEL
D. melanogaster	DGI**K**RICLEL
S. cerevisiae	STL**K**KLSFEL
A. thaliana	PTV**K**KVSLEL

Figure 5.5: Multiple-species protein alignment of the affected amino acid p.K301 within the SSADH protein. The affected amino acid (bold and underlined) is shown together with flanking sequences and marked by an arrow.

[4] (see section 3.3)

5.4 Succinic semialdehyde dehydrogenase deficiency

The clinical phenotype of SSADH deficiency has a high intra- and interfamilial variability ranging from mild delayed intellectual, motor, speech and language development to severe neurological defects including seizures, hypotonia, ataxia and behavioural problems [Jakobs *et al.*, 1993; Gibson *et al.*, 1997; Pearl *et al.*, 2003]. In affected individuals, the GABA degradation pathway is disrupted. GABA, the major inhibitory neurotransmitter of the brain, is derived from the major excitatory neurotransmitter glutamate. After reuptake from the synaptic cleft, GABA is deaminated to succinic semialdehyde (SSA). Subsequently, the NAD^+-dependent SSADH oxidizes SSA to succinate (Equation 5.1, see also Figure 6.1).

$$\mathrm{SSA} + \mathrm{NAD^+} + \mathrm{H_2O} \rightleftharpoons \mathrm{succinate} + \mathrm{NADH} + 2\mathrm{H^+} \tag{5.1}$$

Loss of SSADH activity leads to accumulation of SSA, which is degraded to gamma-hydroxybutyric acid (GHB) [Lyon *et al.*, 2007] that in turn has effects on multiple neurotransmitter systems [Gibson *et al.*, 2003]. The accumulation of GHB in physiological fluids of the affected individuals with concentrations range from 2-fold to 800-fold in urine, 4-fold to 200-fold in plasma, and 100-fold to 1200-fold in cerebrospinal fluid (CSF) compared to controls, which is the biochemical hallmark of SSADH deficiency [Gibson *et al.*, 1990]. As reviewed by Kim *et al.*, SSADH deficiency is usually diagnosed by analysis of physiological fluids, and suggestive GHB levels are then confirmed by enzymatic and molecular tests [Kim *et al.*, 2011].

5.5 p.K301E abolishes SSADH activity

Unfortunately, blood, liquor or urine of affected individuals from family 8600485 were not available. Therefore, SSADH activity was investigated using protein preparations from patient lymphoblasts. During oxidation of SSA to succinate by SSADH, the cofactor NAD^+ is converted to NADH, which fluoresces and has an emission peak at 460 nm. Therefore, NADH fluorescence can be used as a measure of SSADH activity. SSADH activity was determined using lymphoblast preparations from six healthy age- and sex-matched controls and the index patient (V-11; Figure 3.2). Various concentrations of the cofactor NAD^+ (0–0.1 mM) were used. Measurement of SSADH activity was performed in triplicate for each NAD^+ concentration on two consecutive days, respectively. The results of this experiment are shown in Figure 5.6.

In the control cell lines, SSADH activity increased in correlation with elevated amounts of NAD^+. However, in the patient cell line no SSADH activity could be detected. To understand whether SSADH p.K301E requires higher NAD^+ concentrations to oxidate SSA to succinate the assay was further performed with NAD^+ concentrations between 0.5 and 2 mM. However, SSADH was found to be still inactive in the patient cell line, whereas in the control cell lines a dose dependent decrease in SSADH activity was observed (data not shown). Day-to-day variability in this study was in line with assay variability published by others [e. g. Gibson *et al.* 1991].

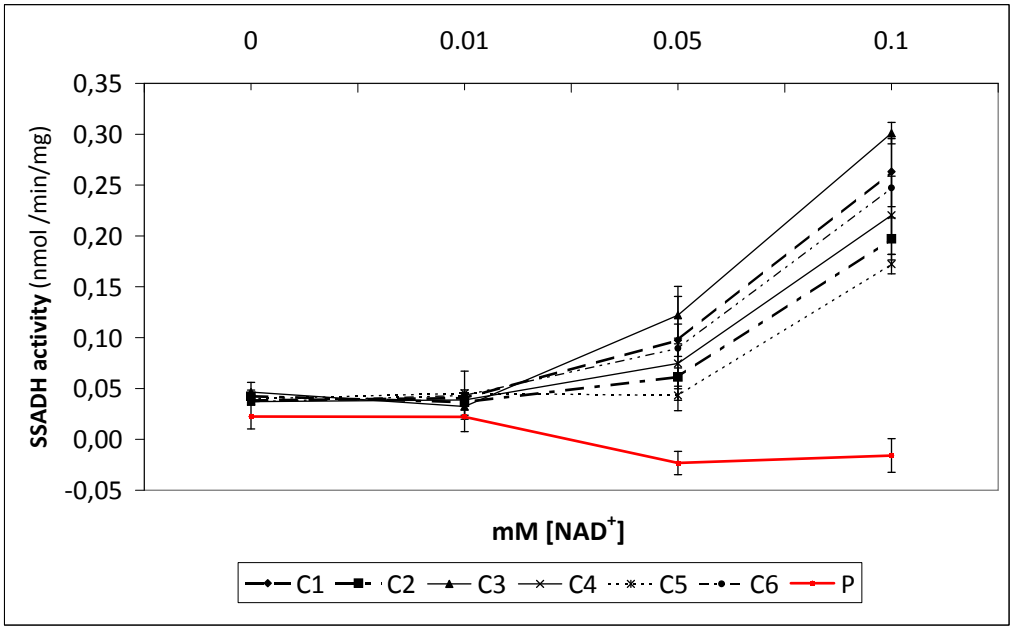

FIGURE 5.6: SSADH activity in lymphoblast cell lines of the affected individual (P) and six controls (C1–C6) at different NAD^+ concentrations (0–0.1 mM). SSADH activity is given as nmol/min/mg protein. Data were normalized against background acivity. Each data point is the mean value of two separate experiments that were performed in triplicate, respectively. Error bars represent the standard deviation.

5.6 Structure-based mutation analysis of SSADH p.K301E

Thus, *in vitro* enzyme tests showed that SSADH p.K301E is enzymatically inactive. To understand the impact of p.K301E on protein structure, *in silico* structure-based modeling of SSADH was performed with the support of Dr. H. Stehr. Modeling of wild-type and mutant SSADH residues was performed based on SSADH crystal structure 2W8R (PDB) (Figure 5.7).

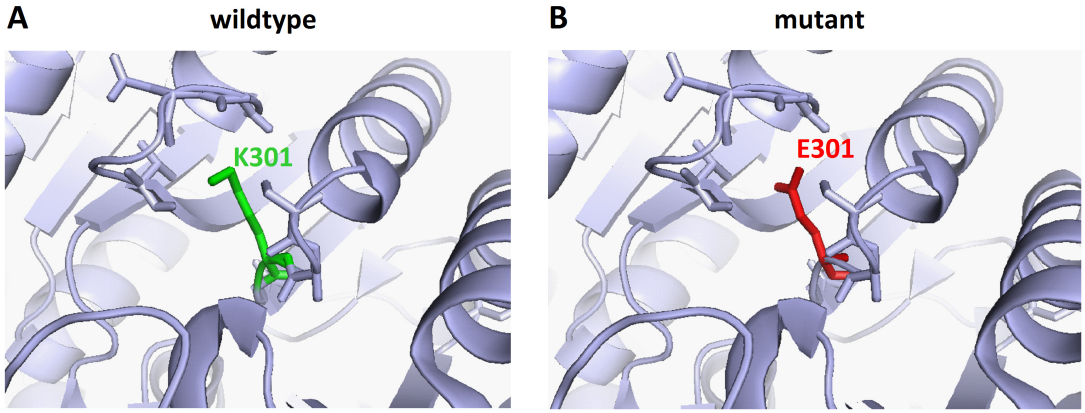

FIGURE 5.7: *In silico* modeling of p.K301E within SSADH. The SSADH residue affected by the missense mutation is presented as stick model in green in the wild-type **(A)** and red in the mutant protein **(B)**. Modeling was perfomed with PDB file 2W8R.

In the wild-type structure, p.K301 connects two α helices and one β strand, which are involved in the binding of NAD^+. The non-conservative change from a positively charged lysine to the negatively charged glutamine likely destabilizes this area of the protein core. Thus, p.K301E SSADH is probably unable to bind NAD^+ and is therefore unable to catalyze the oxidation of SSA to succinate. Dr. H. Stehr quantified this effect by *in silico* energy calculations with the Concoord/PBSA method [Benedix *et al.*, 2009] which calculates mutation-dependent stability changes. The simulation predicted a decrease in stability of 6.67 kcal/mol, which indicates a profound destabilization of the SSADH protein.

These results are in keeping with the abolished SSADH activity in the lymphoblastoid cell line of the affected family member and strongly suggest that the SSADH deficiency in family 8600485 is caused by c.901A > G (p.K301E) in the *ALDH5A1* gene.

6 Discussion

I report here on two consanguineous Iranian families affected by two different forms of ARID. In the first family (M289) the affected individuals suffer from NS-ARID and present with moderate ID and borderline microcephaly without any other co-morbidities. In the second family (M8600485), the affected individuals present with a profound syndromic phenoytpe including moderate ID, tonic-clonic seizures and developmental delay. Using homozygosity mapping and Sanger sequencing, plausible ARID gene defects were identified for both families. Subsequently, functional studies were performed to elucidate the genotype-phenoytpe correlation in both disorders. The findings concerning family M289 are discussed below. The studies of 8600485 are discussed in section 6.2.

6.1 Discussion – Family M289

I present two homozygous missense mutations in two novel NS-ARID candidate genes, *SARS* and *ZNF697*, that segregate with moderate ID and borderline microcephaly in a large consanguineous Iranian family (M289) with arab ethnicity. Both genes are located in a genomic ARID locus on chromosome 1p13.3–p11.2. In this study, I characterized both genes and mutations with respect to their functional implications in cognition and their influence on the ID phenotype observed in the affected individuals. The sequence variant within *ZNF697* (c.472C > A [p.P158T]) is predicted to be benign by four algorithms and the affected nucleotide is not conserved. Therefore, it is likely that the *ZNF697* variant is a previously undescribed rare polymorphism, and the only remaining sequence alteration with disease-causing potential is the c.514G > A mutation in the *SARS* gene.

6.1.1 SARS p.D172N probably underlies ARID in M289

ARSs are essential and ubiquitously expressed enzymes responsible for ligating amino acids to cognate tRNA molecules in mitochondria and in the cytosol. Mutations in five mitochondrial ARS have been associated with brain specific phenotypes and four genes encoding cytoplasmic ARSs have been implicated in inherited peripheral neuropathy with an axonal pathology indicating an important role of this enzyme class in neurons and the human brain [Yao and Fox, 2013; Antonellis and Green, 2008]. Therefore, *SARS* is the most plausible candidate gene for the ARID phenotype observed in family M289. Moreover, the mutation within *SARS* (c.514G > A [p.D172N]) affects a highly conserved nucleotide position and was predicted to have deleterious

consequences on SARS protein function by five different algorithms. Interestingly, a heterozygous nonsense mutation (p.R107X) within *SARS* is reported in the Exome Variant Server but as this was supposedly found in a healthy individual, one could assume that SARS p.R107X transcripts probably undergo degradation mediated by nonsense-mediated mRNA decay (NMD). Thus, functional loss of one SARS allele does not lead to haploinsufficiency in humans. Findings in mice support this assumption, as Seburn *et al.* demonstrated that animals heterozygous for a glycyl-tRNA synthease (GARS) null allele ($Gars^{-/+}$) did not show phenotypic alterations in spite of a twofold lower Gars mRNA level and threefold lower Gars enzyme activity than wild-type mice in brain and kidney [Seburn *et al.*, 2006]. Moreover, also in *D. melanogaster*, loss of one *YARS* allele does not result in haploinsufficieny Storkebaum2009. Therefore, in keeping with the autosomal recessive mode of inheritance observed in M289, loss of one gene copy of an essential ARS gene is apparently not sufficient to cause a disorder.

6.1.2 *SARS* is expressed in brain regions that are important for learning and memory

Cells destined to form the nervous system begin to develop three weeks after fertilization. During the development of the brain and nervous system the emerging neurites need to cover great distances on their way to their target cells. The tip of a neurite is made of the axonal growth cone that finds its path through the central nervous system by following signals sent by cell surfaces, growth factors, growth cone attractants, and growth cone repellents [Tessier-Lavigne and Goodman, 1996; O'Donnell *et al.*, 2009]. Upon reaching its final target, the growth cone is replaced by a presynaptic nerve terminal. In fetal rat brain, a great diversity of proteins is involved in axonal pathfinding, cytoskeletal remodeling, vesicular traffic and carbohydrate metabolism. The majority of these proteins play roles in translation (including SARS and further 18 cytoplasmic and bifunctional ARSs), protein folding, posttranslational processing, and proteasome/ubiquitination-dependent degradation [Estrada-Bernal *et al.*, 2012].

In the study presented here, *SARS* was found to be already expressed during these early stages of brain development but also in the mature brain. RT-PCR experiments using RNA extracted from fetal brain tissue showed *SARS* expression in whole fetal brain as well as in the parietal lobe, the frontal lobe and the temporal lobe. These three parts of the human cerebrum are exceptionally important for memory and learning. Parietal lobe function has been linked to working memory, spatial orientation [Andersen and Buneo, 2002; Cabeza and Nyberg, 2000] and attentional processes [Corbetta and Shulman, 2002]. Higher-level cognitive processes, e. g. executive skills (decision-making, planning, sustained attention, awareness and insight) and working memory, take place in the frontal lobe [Linden, 2007]. The most anterior part of the frontal lobe, the prefrontal cortex, controls many higher-order executive tasks. It is involved in e. g. learning [Pasupathy and Miller, 2005; Antzoulatos and Miller, 2011], memory [Warden and Miller, 2010], categorization [Freedman *et al.*, 2001; Antzoulatos and Miller, 2011] as well as cognitive flexibility [Clarke *et al.*, 2004; Gruber *et al.*, 2010; Rygula *et al.*, 2010]. The medial temporal lobe includes a system of anatomically related structures that

underly conscious memory for facts and events (declarative memory). The system consists of the hippocampal region and the adjacent cortices (perirhinal, entorhinal, and parahippocampal cortex) [Squire *et al.*, 2004, 2007].

In the mature brain expression of *SARS* was observed in extracts from adult hippocampus tissue. The hippocampal axons form two-way connections among the temporal, frontal, and parietal lobes [Robertson, 2002]. Moreover, the hippocampus and its surrounding tissue transfer explicit information to permanent storage sites located throughout the cerebral cortex [Eichenbaum, 1999]. Thus, the hippocampus is the central processing area for declarative learning and memory [Squire, 2004].

Interestingly, Loerch *et al.* found *SARS* to be one of those genes that are differentially regulated in the cortex of humans, rhesus macaques and mice[1] . It was observed that humans and rhesus macaques diverge from mice due to a profound increase in age-dependent depression of neuronal genes [Loerch *et al.*, 2008]. The depression is caused by increased DNA damage in the promoter regions particularly of those genes which are involved in functions relevant for learning, memory and neuronal survival [Lu *et al.*, 2004]. Down-regulation of SARS expression in the aging human cortex therefore underscores the importance for human *SARS* in cognitive processes and its likelihood to be a plausible ARID candidate gene[2].

6.1.3 Confocal microscopy reveals cytoplasmic localization of wild-type and mutant SARS in SH-SY5Y, HeLa and HEK293-T cells

Several mutations within *GARS* and *YARS* that underly peripheral neuropathy cause altered intracellular localization of their specific gene products. For example, in neuronal and non-neuronal cells, GARS tagged with enhanced green fluorescent protein (EGFP) associates with cytoplasmic granules, whereas mutant forms of GARS display a diffuse cellular localization pattern [Antonellis *et al.*, 2006]. A similar observation was made for YARS, which is associated with granular structures in the neurite projections of cultured neurons and in primary embryonic motor neurons. Whereas wild-type YARS tagged with EGFP co-localizes with cytoplasmic granule structures, EGFP-tagged mutant forms of YARS showed marked reduction in granule localization and a diffuse cytoplasmic staining [Jordanova *et al.*, 2006].

To understand whether altered subcellular localization of SARS might be involved in the disease mechanism underlying ID in family M289, expression of endogenous *SARS* and N- or C-terminally tagged *SARS* variants (wild-type and p.D172N) was studied in SH-SY5Y cells as a neuronal model system. In addition, HeLa and HEK293-T cells were investigated to explore potential cell-specific differences. Fluorescence stained endogenous SARS and EYFP- and mCherry tagged forms of SARS (wild-type and p.D172N) localized unanimously in the cytoplasm and no cell specific differences were observed. These results are in good agreement with a study

[1] While in the aging brains of rhesus macaques and mice the expression of *SARS* is up-regulated, *SARS* expression is down-regulated in the human brain [Loerch *et al.*, 2008].

[2] Of note, *ZNF697* expression changes with age in all three species, but differs directionally between mouse and rhesus macaque. In the aging brain of human and rhesus macaque the expression of *ZNF697* is downregulated, but up-regulated in the aging mouse brain [Loerch *et al.*, 2008].

by Xu *et al.* [2012]: using confocal microscopy these authors show that in HUVEC (Human Umbilical Vein Endothelial Cells) the majority of SARS is located in the cytoplasm and only a minority of SARS is present in the nucleus [Xu *et al.*, 2012]. The absence of SARS in the nucleus of SH-SY5Y, HEK293-T and HeLa cells observed here can be explained by the resolution limit of confocal microscopy that might not allow the detection of extremely low amounts of flourescent protein. Furthermore, the presence of high amounts of SARS in the nucleus is probably not needed in these cell types. In contrast, nuclear SARS is essential for angiogenesis and therefore mandatory in HUVEC [Xu *et al.*, 2012]. Specifically, SARS[3] regulates the expression of the vascular endothelian growth factor (VEGFA), a key regulator of vascular development, most likely at the level of transcription [Xu *et al.*, 2012; Herzog *et al.*, 2009; Fukui *et al.*, 2009].

Still, as the mutation p.D172N within SARS apparently did not lead to considerable mislocalization detectable by confocal microscopy, I set out to further investigate by cell fractionation, whether the EYFP-tagged wild-type SARS and p.D172N SARS could be detected in the nucleus of mammalian cells and if the p.D172N substitution has an influence on subcellular SARS concentrations.

6.1.4 Expression analysis reveals cytoplasmic and nuclear localization of wild-type and mutant SARS in mammalian cells

Energy calculations of stability changes upon mutation revealed a destabilizing effect of the p.D172N substitution on SARS. Concordant with this *in silico* prediction, expression analysis suggested that the mutant protein could be unstable. In fact, expression of ectopic SARS p.D172N was highly significantly reduced as compared to ectopic wild-type SARS in whole cell lysate. Moreover, less ectopic mutant protein was also detected in the cytosolic compartment, confirming our previous results. To exclude differences attributed to unequal transfection efficiency, immunofluorescence experiments were performed. Transfection rates for both expression vectors (SARS wild-type and p.D172N) were found to be equally 40 %.

Furthermore, these experiments revealed that overexpressed SARS wild-type and mutant proteins are not only present in the cytosol but also in the nuclear fraction. These findings were also confirmed for endogenous SARS (see section 9.4). In the nuclear fraction equal amount of ectopic EYFP-SARS proteins were observed, suggesting that p.D172N does not affect the translocation in this compartment. However, quantification was hampered by the low protein amount. Although the nuclear function of most ARSs is not yet clear, it has been proposed that ARSs might perform aminoacylation-dependent quality control of tRNA to ensure that only fully mature tRNA is exported into the cytoplasm [Grosshans *et al.*, 2000; Sarkar *et al.*, 1999; Lund and Dahlberg, 1998].

[3] SARS is one of four ARSs (EPRS [Jia *et al.*, 2008; Ray and Fox, 2007], YARS [Wakasugi *et al.*, 2002a] and WARS [Wakasugi *et al.*, 2002b]) that regulate angiogenesis by a non-canonical activity that is independent of their aminoacylation function.

6.1.5 p.D172N within SARS impairs aminoacylation activity

In silico modeling revealed that the p.D172N substitution might affect the aminoacylation activity of SARS, because p.D172N is located spatially close to the active site, but faces away from the seven-stranded antiparallel β-sheet (β1–β9–β10–β11–β13–β8–β7) that constitutes the central core of the aminoacylation domain. However, p.D172 provides a negative charge that is lost in the SARS p.D172N variant. This loss might have negative effects on the hydrophobic pocket next to the active site. To test this hypothesis, I purified recombinant SARS proteins (wild-type, p.D172N and p.T429A) and investigated the effect of the missense mutations on the first step of the aminoacylation reaction, where serine is activated and seryl-adenylate is formed.

The activity of SARS wild-type, SARS p.D172N and the aminoacylation-defective SARS protein (p.T429A) was assayed for 90 minutes *in vitro*. Remarkably, this analysis indicated that the p.D172N substition signifcantly impairs the ability of SARS to activate serine. The amount of pyrophosphate released during this reaction was significantly lower as compared to the amount of pyrophosphate released by the SARS wild-type enzyme, when incubated for 60 minutes. Even after 90 minutes, SARS p.D172N released 51 % less pyrophosphate than the wild-type enzyme after 60 minutes.

These findings are in good agreement with reduced aminoacylation activity of a previously described SARS p.F383V variant [Xu *et al.*, 2012]. The residue p.F383 is located near the end (β10) of the same core β-strand and near the active site. The side chain of p.F383 forms hydrophobic interactions with p.H170 and p.F316 to stabilize the β10–β11 hairpin. Analogous to the *in silico* model for SARS p.D172N, the close proximity to the active site and the disturbance of the hydrophobic core next to the active site are assumed to cause the reduced aminoacylation activity of SARS p.F383V [Xu *et al.*, 2012]. Together, these data suggest that p.D172N impairs SARS activity by indirect disturbance of the central core of the aminoacylation domain and thus, that p.D172 is indispensible for normal aminoacylation acivity of SARS.

The highest amount of pyrophosphate released by the SARS wild-type enzyme was observed after 60 minutes of incubation, however, the amount of pyrophosphate was reduced down to 67 % at 90 minutes. This decrease in SARS wild-type serine-activation can be explained by several mechanisms. Firstly, AMP produced during the formation of seryl-adenylate could enter into the corresponding back-reaction followed by γ-phosphate transfer from ATP. It was shown by Rapaport and coworkers that this mechanism is an inherent part of the catalytic activities of ARSs and could prevent the buildup of enzyme-bound or free aminoacyl adenylates [Rapaport *et al.*, 1987]. Secondly, substrate inhibition might cause the decrease in aminoacylation activity because seryl-adenylate can remain bound to the SARS enzyme. During aminoacylation of *E. coli* alanyl-tRNA synthetase (AARS), the alanyl-adenylate product remains bound tightly to the enzyme and is only turned-over when $tRNA^{Ala}$ is present in the reaction. In addition, pyrophosphate generated in the adenylation step substantially inhibits the initial rate of aminoacylation of $tRNA^{Ala}$ in a dose dependent manner. The degree to which pyrophosphate inhibits a given reaction depends, inter alia, on the extent to which the reaction has progressed. However, the

mechanism of this inhibition still needs to be established [Wolfson and Uhlenbeck, 2002].

6.1.6 SARS p.D172N probably leads to severely reduced amounts of serine-tRNA^{Ser} and serine-tRNA^{Sec} available for translation

It was shown in this study that *SARS* is well expressed in tissues relevant for memory and learning during fetal brain development and in the mature brain. What is more, the study provides evidence that the missense mutation c.514G > A [p.D172N] within *SARS* does not only lead to destabilization of the SARS protein but also impairs the aminoacylation activity of SARS p.D172N.

Both defects together probably lead to severely reduced amounts of charged tRNA^{Ser} and tRNA^{Sec} available for translation at the ribosome, which in turn could lead to a reduced translation rate, especially of those proteins that contain high amounts of serine or selenocysteine. The latter amino acid is an essential component of the 25 selenoproteins encoded in the human genome [Lu and Holmgren, 2009]. Selenoproteins are involved in protein folding, degradation of misfolded membrane proteins, and control of cellular calcium homeostasis [Andersen, 2004]. Interestingly, deletion of selenoprotein P (*Sepp1*) in mice produces both neuronal and axonal degeneration as well as potentially reversible neurite changes in the developing brain [Caito *et al.*, 2011]. Neuron-specific ablation of selenoprotein expression causes a neurodevelopmental and -degenerative phenotype in mice, affecting the cerebral cortex and the hippocampus, particularly the parvalbumin (PV) -positive interneuron population [Wirth *et al.*, 2010]. Moreover, loss of PV neurons is associated with some neuropsychiatric disorders, e. g. autism spectrum disorders [Sgado *et al.*, 2013].

Furthermore, SARS p.D172N probably slows down overall translation rates during cellular processes that depend on fast and reliable protein translation (e. g. in axonal growth cones (see above and reviewed by e. g. [Jung and Holt, 2011; Swanger and Bassell, 2011]) or synaptic processes involved in learning, memory processing and memory storage as reviewed by e. g. [Gal-Ben-Ari *et al.*, 2012; Bekinschtein *et al.*, 2010; Costa-Mattioli *et al.*, 2009]). Similar disease mechanisms have been proposed for mutations within *LARS*, underlying infantile hepatopathy [Casey *et al.*, 2012], and mutations within *KARS* that cause recessive intermediate Charcot-Marie-Tooth disease (CMT) type B (CMTRIB) (MIM #613641) [McLaughlin *et al.*, 2010]. Analogous to the homozygous missense mutation in *SARS*, a homozygous missense change in *LARS* is suggested to cause inefficient aminoacylation of tRNA^{Leu} and to result in reduced rate of protein synthesis in the liver. This impaired translation may in turn affect specifically synthesis of proteins with a high leucine content [Casey *et al.*, 2012]. In the case of *KARS*, compound heterozygosity for a missense mutation that severely reduces KARS aminoacylation activity and a 2-bp insertion resulting in a null allele have been identified. Interestingly, the CMTRIB phenotype does not only comprise peripheral neuropathy but also self-abusive behavior and developmental delay [McLaughlin *et al.*, 2010].

SARS seems to belong to the many ubiquitously expressed genes associated with ID with indispensable cellular functions, such as DNA transcription and translation, protein degradation and

mRNA splicing. Concordant with the mode of inheritance and phenotype observed in family M289, most of these defects are associated with NS-ARID. Why clinical consequences of mutations in essential genes are restricted to the brain is still unclear. However, synapses might be particularly vulnerable to imbalances in the cell- or energy metabolism [Najmabadi *et al.*, 2011]. As reviewed by [Valnegri *et al.*, 2012] more than 50 % of ID-related proteins are enriched in synpatic compartments and may be involved in synaptic plasticity, synapse formation and cytoskeleton rearrangement [see also Ropers and Hamel, 2005]. The hypothesis that some features of ID are caused by alterations in synaptic functions is underscored by hystological data, which further show a correlation between the severity of ID and the severity of dendritic spine malformations [Valnegri *et al.*, 2012]. Furthermore, higher brain functions have a complex molecular basis and specifically the cognitive abilities of humans represent have only recently evolved [Hu *et al.*, 2011].

SARS is the first ARSs gene to be associated with ARID. In the course of our systematic clinical studies and autozygosity mapping in large consanguineous Iranian families with ARID, we have recently identified mutations within a mitochondrial ARSs gene segregating with SID in one family (data not shown). This finding underscores the relevance of SARS for the cognitive phenotype observed in family M289. However, further in depth functional investigations *in vitro* and *in vivo* are needed to shed more light on the precise role of SARS during brain development and in differentiated neuronal tissues.

6.2 Discussion – Family M8600485

ID is not a single condition, but the phenotypic hallmark of a collection of syndromic and non-syndromic disorders. The majority of ID cases still remains unexplained [Rauch *et al.*, 2006]. Even if an affected individual presents with a syndromic phenotype, giving a definite diagnosis is all but impossible, since overlapping clinical presentations could be caused by different genetic loci.

The affected individuals of family M8600485 presented with a complex phenotype encompassing developmental delay as well as speech delay, severe ID, tonic-clonic seizures, and hypotonia during the neonatal period. In order to identify the underlying genetic defect, genotyping (SNP analysis) and linkage analysis were carried out and a linkage interval on chromosome 6 was identified, containing more than 280 genes. Interestingly this interval contained *ALDH5A1*, defects in which, are known to been associated with congenital SSADH deficiency.

SSADH deficiency is a rare autosomal recessive disorder and parental consanguinity is observed in approximately 40 % of the cases [Pearl *et al.*, 2003]. The SSADH deficient patient database maintained by the Department of Neurology at Children's Medical Center in Washington DC is based on systematic questionnaire data of 60 patients. These data suggest that developmental delay as well as ID are global findings in affected individuals (100 %), and that hypotonia (82 %), ataxia (77 %) as well as tonic-clonic seizures (53 %) are common clinical symptoms [Pearl *et al.*, 2009, 2011]. All of these features, except for ataxia, are present in the affected individuals in our study and are consistent with the typical characteristics of human SSADH deficiency. Therefore, in this study, the coding region of *ALDH5A1* was sequenced, which led to the identification of a previously undiscribed missense change (c.901A > G, leading to SSADH p.K301E). Furthermore, I could show that this mutation leads to undetectable levels of SSADH activity.

6.2.1 Influence of defective SSADH on GABA metabolism

In affected individuals, the degradation pathway of GABA (gamma-aminobutyric acid), the major inhibitory neurotransmitter of the brain, is disrupted (Figure 6.1). In SSADH deficiency oxidation of SSA to succinate is reduced or abolished and succinic semialdehyde reductase (aflatoxin B1 aldehyde reductase member 2, AKR7A2) uses SSA as a substrate to produce gamma-hydroxybutyric acid (GHB) [Maitre, 1997; Kelly *et al.*, 2002; Gibson *et al.*, 2003; Gropman, 2003]. The neurotransmitter GHB is a short fatty acid that occurs naturally in the mammalian brain and is normally found at a level < 1 % of GABA [Doherty *et al.*, 1978]. GHB has effects on multiple neurotransmitter systems, such as dopamine, serotonine, acetylcholine and GABA [Gibson *et al.*, 2003]. Exogenously given GHB crosses the blood-brain barrier and leads to unusual behaviour as well as abnormal electrophysiological and biochemical dose dependent effects [Snead, 2000; Snead and Gibson, 2005; Wong *et al.*, 2004a,b]. These effects can comprise short-term amnesia and memory loss (low dosage), seizures and sleep induction (moderate dosage), stupor, coma, and potentially respiratory arrest at very high dosages [Snead and Gibson, 2005].

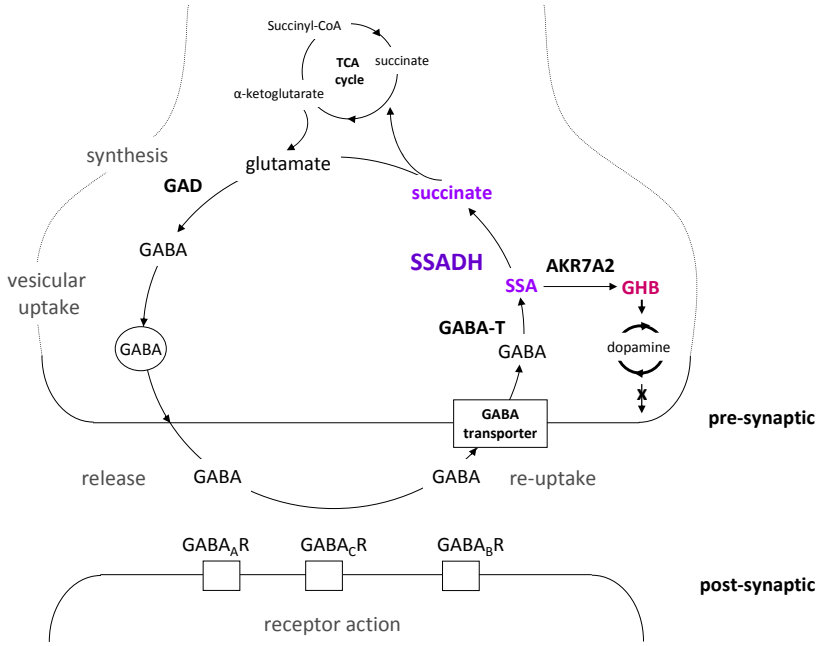

FIGURE 6.1: GABA metabolism at the synapse. The tricarboxylic acid (TCA) cycle is one source of glutamate. Glutamate is decarboxylased by glutamic decarboxylase (GAD) to GABA. GABA is packaged into vesicles and released into the synaptic cleft, where it can bind to GABA receptors. GABA is removed from the synaptic cleft by GABA transporters (GAT) and is subsequently converted to SSA by GABA-transaminase (GABA-T). SSA is normally oxidized to succinate, which in turn enters the TCA cycle. In SSADH deficiency, SSA accumulates and can be converted to GHB by aldo-keto reductase 7A2 (AKR7A2). At increased levels, GHB inhibits presynatic release of dopamine and enhances dopamine turnover. Illustration according to [KIM *ET AL.*, 2011].

In SSADH deficiency, GHB is massively over-produced and accumulates in blood, urine or cerebrospinal fluid (CSF). In the CSF of affected individuals levels of GHB and GABA are elevated (65- to 230-fold and up to 3-fold, respectively), and levels of glutamine are decreased [Gibson *et al.*, 2003]. Concordant with these findings, SSADH deficient mice (*Aldh5a1*$^{-/-}$) show a 60-fold increase of GHB and a 2-fold increase of GABA levels, as well as decreased glutamine levels in the brain [Jansen *et al.*, 2008; Hogema *et al.*, 2001]. In physiological concentrations, GHB only binds at the GHB receptor [Snead, 2000; Wu *et al.*, 2004], however, in high concentrations, such as those observed in the brains of patients with SSADH deficiency, GHB also binds to $GABA_B$ receptors [Gervasi *et al.*, 2003; Buzzi *et al.*, 2006]. $GABA_B$ receptors located in the postsynaptic neuron are associated with receptor G-protein-coupled inwardly rectifying potassium channels (GIRKs). High levels of GHB and (and GABA) thus lead to a decrease in GIRK channel function, which is believed to be the underlying cause for seizures in human and murine SSADH deficiency [Vardya *et al.*, 2010].

6.2.2 p.K301E prevents NAD$^+$-binding and abolishes SSADH activity

In SSADH deficiency the activity of SSADH is reduced or nearly abolished. To investigate whether the p.K301E substitution has an impact on SSADH function, I studied SSADH enzyme activity in the lymphoblastoid cell line of one affected individual and six unaffected controls. I could show that SSADH activity is completly abolished in the affected individuals of M8600485. In controls, SSADH activity increased with elevated amounts of NAD^+ in the range from 0–0.1 mM NAD^+. Higher NAD^+ concentrations led to a decrease of SSADH activity, however, during oxidation of SSA to succinate, NAD^+ is reduced to NADH, which is known to exhibit an inhibitory effect for several mammalian SSADHs [Duncan and Tipton, 1971; Blaner and Churchich, 1979; Rivett and Tipton, 1981; Kang *et al.*, 2005]. This might explain the inhibition of SSADH activity I observed in the cell lysates from the healthy control individuals at NAD^+ leves > 0.1 mM. What is more, in a cellular context, inhibition of SSADH by NADH leads to accumulation of SSA, which in turn leads to further inhibition of SSADH [Kammeraat and Veldstra, 1968; Blaner and Churchich, 1979].

To understand the molecular mechanism underlying this severe reduction in SSADH activity, structure based *in silico* modeling was employed. The lysine affected by the p.K301E mutation in family 8600485 is located in close spatial proximity to p.Q306 and p.C340, which together constitute the active site environment. Under oxidized conditions SSADH is inactive, a disulfide bond is formed between the catalytic residues p.C340 and p.C342 and a catalytic loop including residues 334–344 blocks the binding sites for SSA as well as for NAD^+ [Kim *et al.*, 2009]. *In silico* analysis provided evidence that p.K301 is part of the protein core and involved in binding of NAD^+. According to these findings, the p.K301E substitution could lead to a severe destabilization of SSADH and prevent binding of NAD^+ to the enzyme. Thus, in the affected individuals of M8600485 SSADH is unable to catalyze the oxidation of SSA to succinate.

This finding is in good agreement with the results of Akaboshi *et al.* who assayed the ac-

tivity of 27 disease causing missense mutations observed in affected individuals. Five missense mutations (p.G176R, p.G268E, p.N335K, p.G409D and p.G533R) showed a nearly abolished enzyme activity (< 1–1 %) [Akaboshi *et al.*, 2003]. While three of these mutations probably affect protein stability (p.G409D) or stability and oligomerization (p.G176R and p.G533R), p.G286E and p.N335K affect the catalytic function of SSADH. The residue p.N335 is located on the "dynamic catalytic loop" and probably leads to a severe distortion of the active site environment or reduced dynamics of the "catalytic loop" or both [Kim *et al.*, 2009]. However, only p.G268 is involved in binding of NAD^+, being one of the residues creating the binding pocket for the adenine base of NAD^+. Thus p.G268E might lead to a loss of NAD^+ binding ability and consequently to a loss of SSADH activity [Kim *et al.*, 2009].

In summary, this study identified a novel homozygous missense mutation in *ALDH5A1* that is associated with SSADH deficiency and severe ID. The mutation is located near the active site and is thus only the second mutation identified to date that might not necessarily cause a loss of SSADH activity solely through protein misfolding and subsequent degradation, but rather affect SSADH activity through an impairment of NAD^+ binding.

6.2.3 Therapeutic options

To date, treatment of SSADH deficiency is generally symptomatic and targeted (therapeutic concepts in human and murine SSADH deficiency have been reviewed e. g. by [Knerr *et al.*, 2007; Kim *et al.*, 2011; Vogel *et al.*, 2012]). Typically, antiepiletics are chosen to relief generalized epilepsy, like in family M8600485, where seizures were successfully treated with the anticonvulsant carbamazepine. The most widely used drug in SSADH deficiency is vigabatrin, an irreversible inhibitor of GABA-transaminase. Vigabatrin is generally used for the treatment of infantile spasms and increases GABA levels in the brain. In the case of SSADH deficiency, the physiologocal role of vigabatrin is to decrease the production of SSA, and as a consequence decrease the amount of GHB [Gropman, 2003]. In SSADH deficient mice, vigabatrin significantly increased the lifespan [Hogema *et al.*, 2001]. However, vigabatrin treatment does not show beneficial results in all SSADH deficiency patients as lack of positive effects and even worsening of symptoms have been reported [Pearl *et al.*, 2009].

The most promising drug today is SGS-742, a GABA analog that acts as an antagonist of the $GABA_B$ receptor and could therefore block supraphysiological GABA levels [Farlow, 2009]. In *Aldh5a1*$^{(-/-)}$ mice, treatment with a progenitor of SGS-742 significantly extended the animals' lifespan [Hogema *et al.*, 2001; Gupta *et al.*, 2002]. Most important, SGS-742 showed positive effects in a Phase II double-blind, placebo-controlled clinical trial in patients with mild cognitive impairment [Froestl *et al.*, 2004]. Furthermore, it was shown that SGS-742 significantly improves the spike-wave duration in a dose dependent manner and also controls absence seizures [Pearl *et al.*, 2009] and future plans are to test the application of SGS-742 in affectd individuals with SSADH deficiency [Kim *et al.*, 2011].

6.3 The future of molecular genetic diagnosis and counseling: comprehensive entry tests

In the diagnosis of genetic disease during the past decades, Sanger sequencing has been regarded as the gold standard for the identification of mutations because it has high accuracy, sensitivity and specificity. The classical strategy for identifying recessive disorders included linkage analysis followed by either systematically sequencing the genes located within the linkage interval (as e. g. for family M289) or applying a candidate gene approach (as e. g. for family 8600485). However, these strategies are costly and time-consuming, and thus not scalable or cost-effective for testing large panels of families/patients [Majewski *et al.*, 2011].

Next generation sequencing (NGS) coupled with bioinformatic processing has become a powerful tool for identifying the underlying gene defects in known and new genes for disorders with genetic heterogeneity (as e. g. ID). Several studies have been published in recent years demonstrating the efficiency of the NGS approach in elucidating the molecular defectcs of several disorders, including ID, see e. g. [Rauch *et al.*, 2012; de Ligt *et al.*, 2012; Najmabadi *et al.*, 2011; Hu *et al.*, 2009]. Targeted NGS is now emerging as the ideal technology for clinical diagnosis. Enrichment targets vary from the exome (whole exome sequencing; WES) to mutation-harboring regions of genes relevant to specific clinical presentations. Recently, a comprehensive test for carrier screening and molecular diagnostic testing was designed, which allows the simultanous screening of 595 genes implicated in recessive diseases [Kingsmore, 2012]. The test includes also a number of genes associated with an ID phenotype, such as e. g. *ALDH5A1* or *ARX* and *SLC6A8*[4]. This test is suggested to increase the rate of successful molecular diagnosis drastically, which will be extremely beneficial for the affected individuals, since to date less than 50 % of patients undergoing serial molecular diagnostic testing receive a definitive diagnosis [Kingsmore, 2012]. Furthermore, such tests will significantly shorten the time required to establish the diagnosis [Ropers, 2012]. Also comprehensive newborn screening for treatable or preventable Mendelian diseases is now being discussed, as it does not only allow early diagnosis but also open up the opportunity for treatment to start even while affected neonates are still asymptomatic for disorders with a later onset (see also e. g. [Saunders *et al.*, 2012]). Timely treatment can also often diminish the clinical severity of conditions and could provide a framework for centralized assessment of investigational new treatments before organ failure [Bell *et al.*, 2011].

Comprehensive entry tests of affected individuals, newborns, couples from risk background[5] or undergoing *in vitro* fertilization are expected to have several beneficial effects: (i) prevention of death and diminished disease severity. (ii) improved quality of life. (iii) narrow the differential diagnosis. (iv) genetic counseling about risks for relatives and in additional offspring. (v) psychosocial benefits. (vi) improved variant database, and (vii) increased understanding of the

[4] (see also subsection 1.1.1)

[5] Populations with a high risk background for recessive disorders include e. g. Arab populations, Amish populations or Ashkenazi Jewish populations. In the latter, for example, preconception testing dramatically reduced the incidence of Tay-Sachs disease (TSD; MIM #272800) [Kingsmore, 2012]

disease mechanism; for a detailed review see [Kingsmore *et al.*, 2011; Kingsmore, 2012].

However, a number of patients may still be left without a definitive diagnosis, as the disease causing mutations might not be in the genes targeted by these tests. Most tests are generally confined to the exonic regions of the relevant genes. Therefore, they will miss a proportion of the disease-causing sequence variants. This challenge can be approached by WES and even in a gloabal setting by whole genome sequencing (WGS), which has the potential to detect almost all genetic variants in the human genome [Ropers, 2012; Lupski *et al.*, 2010; Sobreira *et al.*, 2010]. For example, recently, using a combination of WES and WGS, heterozygous, *de novo* truncating mutations in *ASXL3* were identified as the underlying gene defect in four unrelated patients with a previously undiagnosed Bohring-Opitz-like syndrome [Bainbridge *et al.*, 2013]. Thus, a most plausible scenario for future molecular genetic diagnosis and counseling will eventually be comprehensive entry tests based on WGS [Ropers, 2012].

7 Summary

Intellectual disability (ID) is the hallmark of an extremely heterogeneous group of disorders that comprises a wide variety of syndromic and non-syndromic phenotypes. I report here on two consanguineous Iranian families with members that are affected by autosomal recessive ID and were recruited during a collaborative project between the Max Planck Institute for Molecular Genetics, Berlin, Germany (Prof. H. H. Ropers) and the Genetics Research Center (GRC) at the University of Social Welfare and Rehabilitation Sciences, Tehran, Iran (Prof. H. Najmabadi). Using homozygosity mapping and Sanger sequencing, the most likely gene defects underlying ARID were identified for both families. Subsequently, functional studies were performed to elucidate the genotype-phenoytpe correlation in both disorders.

The affected individuals of family M289 present with moderate ID and borderline microcephaly without any other co-morbidities. The only sequence alteration I identified in this family is a mutation (c.514G > A [p.D172N]) in the *SARS* gene, which encodes the cytoplasmic seryl-tRNA synthetase SARS that is responsible for charging $tRNA^{Ser}$ and $tRNA^{Sec}$ with serine (aminoacylation). *In silico* modeling of SARS revealed that p.D172 is located in proximity to the active site and provides a negative charge, which is lost upon mutation. To understand whether the aminoacylation reaction is affected, the capability of serine activation of SARS wild-type (WT) and mutant proteins was studied in a pyrophosphate (PP_i) release assay. After 60 minutes, SARS p.D172N showed significanty less PP_i release as compared to the WT enzyme and even after 90 minutes the mutant enzyme produced still 51 % less PP_i than the WT at 60 minutes. Moreover, concordant with *in silico* predictions, expression studies comparing WT and mutant *SARS* in mammalian cells provided evidence that SARS p.D172N is unstable.

Reduced aminoacylation and instability of SARS may lead to severely reduced amounts of cognate serine-tRNAs in the cytoplasm, which in turn may cause decreased translation rates. This could have unfavourable consequences for cellular processes that depend on fast and reliable protein translation such as neurite outgrowth or synaptic processes involved in learning and memory. However, further in depth functional investigations are needed to shed more light on the precise role of this protein during brain development and in fully differentiated neuronal tissues.

In family M8600485 affected individuals presented with a profound syndromic phenoytpe including ID, hypotonia and tonic-clonic seizures. In the already available genome wide linkage data from this family, I found *ALDH5A1* to be present within the sole linkage interval with significant LOD-Score on chromosome six. As the patient phenotype was strongly overlapping with

the features commonly observed in individuals affected with congenital succinic semialdehyde dehydrogenase (SSADH) deficiency, which is caused by mutations in *ALDH5A1*, I sequenced this gene in the index patient of M8600485. This revealed a novel missense mutation c.901A > G [p.K301E] within *ALDH5A1* cosegregating with the SSADH deficiency phenotype.

In silico protein modeling showed that p.K301E leads to a putative loss of NAD^+ binding. Concordant with this prediction, no SSADH enzyme activity could be detected in patient lymphoblasts at various NAD^+ concentrations. The mutation p.K301E is located near the active site and is only the second mutation identified to date that might affect SSADH activity through an impairment of NAD^+ binding.

Thus, the gene defects underlying the respective ARID phenoytpes could be identified for both families. This conclusion was further corroborated by the *in silico* analyses and *in vitro* experiments also carried out in this study, which provided additional evidence as to the functional impact of the mutations on important cellular processes.

8 Zusammenfassung

Besonders autosomal-rezessive mentale Retardierung (ARMR) ist durch ausgeprägte Heterogenität gekennzeichnet und umfasst eine große Vielfalt von syndromalen und nicht-syndromalen Formen. Im Zuge eines übergreifenden Kooperationsprojekts zwischen dem Max-Planck-Institut für Molekulare Genetik, Berlin (Prof. H. H. Ropers) und dem Genetics Research Center (GRC) der University of Social Welfare and Rehabilitation Sciences, Tehran, Iran (Prof. H. Najmabadi) zur Identifizierung der genetischen Ursachen von ARMR, wurden im Rahmen dieser Arbeit zwei iranische, konsanguine Familien mit mental retardierten Mitgliedern untersucht. Mittels Autozygotie-Kartierung und anschließender Sanger-Sequenzierung wurden die höchstwahrscheinlich ursächlichen Gendefekte identifiziert. Beide Gendefekte kosegregierten mit dem Phänotyp und wurden nicht in Kontrollen gefunden.

Die Patienten der Familie M289 weisen eine mittelschwere Form mentaler Retardierung auf und haben neben grenzwertiger Mikrozephalie keine weiteren Begleitsymptome. Als einzige Mutation mit krankheitsauslösendem Potential wurde die c.514G > A [p.D172N] Sequenzvariante in *SARS* identifiziert. *SARS* kodiert die cytoplasmatische Seryl-tRNA-Synthetase, deren Funktion die Beladung von $tRNA^{Ser}$ und $tRNA^{Sec}$ mit Serin ist (tRNA-Aminoacylierung).

Anhand von *in silico*-Protein-Modellierung wurde gezeigt, dass p.D172 in der Nähe des aktiven Zentrums von SARS liegt und mit einer negativen Ladung wahrscheinlich zur Aufrechterhaltung einer hydrophoben Umgebung beiträgt. Diese Ladung geht durch die p.D172N-Substitution verloren. Um herauszufinden, ob die p.D172N-Mutation die enzymatische Funktion von SARS beeinträchtigt, wurde die Freisetzung von Pyrophosphat (PP_i) während des ersten Reaktionsschritts der tRNA-Aminoacylierung untersucht. Nach 60 Minuten wurde von SARS p.D172N hoch signifikant weniger PP_i freigesetzt als von dem SARS Wildtyp-Enzym (WT) und auch nach 90 Minuten betrug die freigesetzte PP_i-Menge 51 % weniger im Vergleich zur PP_i-WT-Menge nach 60 Minuten. Weiterhin wurde mittels Western Blot und Zellfraktionierung gezeigt, dass SARS p.D172N in menschlichen Zellen wahrscheinlich instabil ist.

Durch Instabilität und reduzierte Enzymaktivität des SARS-p.D172N-Proteins steht dem Translationsprozess wahrscheinlich eine reduzierte Serin-$tRNA^{Ser/Sec}$-Menge zur Verfügung. Dies könnte die Geschwindigkeit der Translation reduzieren und zu Defiziten bei Vorgängen führen, welche eine schnelle und zuverlässige Proteinsynthese benötigen (beispielweise die Bildung synaptischer Verbindungen wie sie bei Lernprozessen im Gehirn stattfinden). Weiterführende funktionelle Experimente in Neuronen und im differenzierten neuronalen Gewebe können dazu beitragen die genaue Funktion von SARS im Rahmen kognitiver Prozesse aufzuklären.

Die Patienten der Familie M8600485 haben ein syndromales Krankheitsbild mit schwerer mentaler Retardierung, Entwicklungsverzögerung, verzögerter Sprachentwicklung und generalisierten tonisch-klonischen Anfällen. Anhand der bereits vorliegenden Ergebnissen der Autozygotie-Kartierung konnte ich in einem Kopplungsintervall mit signifikantem LOD-Score das Gen *ALDH5A1* identifizieren. Da das Krankheitsbild der Patienten typisch für Succinat-Semialdehyd-Dehydrogense (SSADH)-Defienz ist, welche durch Mutationen in *ALDH5A1* verursacht wird, wurde dieses Gen im Indexpatienten sequenziert. Hierbei wurde die bisher nicht bekannte Missense-Mutation c.901A > G [p.K301E] in *ALDH5A1* identifiziert. *In silico*-Protein-Modellierung dieser Mutation zeigte, dass p.K301E vermutlich zum Verlust der Fähigkeit NAD^+ zu binden führt. Zusätzlich konnte berechnet werden, dass die Mutation eine deutliche Abnahme der freien Energie verursacht und damit zu einer starken Destabilisierung von SSADH führen kann. Übereinstimmend mit diesen Vorhersagen wurde der Verlust der SSADH-Enzymaktivität bei unterschiedlichen NAD^+-Konzentrationen in Proteinextrakten aus Lymphoblasten eines Patienten aus Familie 8600485 nachgewiesen. Die Mutation p.K301E ist erst die zweite Missense-Mutation, welche die Bindung von NAD^+ an SSADH beeinträchtigt und zu nicht nachweisbarer SSADH-Aktivität führt.

In dieser Arbeit konnten somit die für die mentale Retardierung ursächlichen Gendefekte in beiden Familien identifiziert werden. Des Weiteren wurden mittels *in silico*-Studien und durch *in vitro*-Experimente die molekularen Hintergründe untersucht und über die so erzielten Ergebnisse ein Zusammenhang mit den Krankheitsbildern und den zu Grunde liegenden molekularen Prozessen hergestellt.

Bibliography

Abbasi-Moheb L, Mertel S, Gonsior M, Nouri-Vahid L, Kahrizi K, Cirak S, Wieczorek D, Motazacker MM, Esmaeeli-Nieh S, Cremer K, Weissmann R, Tzschach A, Garshasbi M, Abedini SS, Najmabadi H *et al.* (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* 90, 847–855.

Abecasis G, Altshuler D, Auton A, Brooks L, Durbin R, Gibbs R, Hurles M, and McVean G (2010) A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073.

Abou Jamra R, Wohlfart S, Zweier M, Uebe S, Priebe L, Ekici A, Giesebrecht S, Abboud A, Al Khateeb MA, Fakher M, Hamdan S, Ismael A, Muhammad S, Nöthen MM, Schumacher J *et al.* (2011) Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity. *Eur. J. Hum. Genet.* 19, 1161–1166.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, and Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249.

Akaboshi S, Hogema BM, Novelletto A, Malaspina P, Salomons GS, Maropoulos GD, Jakobs C, Grompe M, and Gibson KM (2003) Mutational spectrum of the succinate semialdehyde dehydrogenase (ALDH5A1) gene and functional analysis of 27 novel disease-causing mutations in patients with SSADH deficiency. *Hum. Mutat.* 22, 442–450.

Andersen J (2004) Oxidative stress in neurodegeneration: cause or consequence? *Nat. Med.* 10, S18–25.

Andersen R and Buneo C (2002) Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189–220.

Antonellis A and Green ED (2008) The role of aminoacyl-tRNA synthetases in genetic diseases. *Annu. Rev. Genomics. Hum. Genet.* 9, 87–107.

Antonellis A, Lee-Lin S, Wasterlain A, Leo P, Quezado M, Goldfarb L, Myung K, Burgess S, Fischbeck K, and Green E (2006) Functional analyses of glycyl-tRNA synthetase mutations suggest a key role for tRNA-charging enzymes in peripheral axons. *J. Neurosci.* 26, 10 397–10 406.

Antzoulatos E and Miller E (2011) Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron.* 71, 243–249.

Arnez JG and Moras D (1997) Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.* 22, 211–216.

Bainbridge M, Hu H, Muzny D, Musante L, Lupski J, Graham B, Chen W, Gripp K, Jenny K, Wienker T, Yang Y, Sutton V, Gibbs R, and Ropers H (2013) De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. *Genome Med.* 5, 11 [Epub ahead of print].

Eichen baum H (1999) The hippocampus and mechanisms of declarative memory. *Behav. Brain. Res.* 103, 123–133.

Bekinschtein P, Katche C, Slipczuk L, Gonzalez C, Dorman G, Cammarota M, Izquierdo I, and Medina J (2010) Persistence of long-term memory storage: new insights into its molecular signatures in the hippocampus and related structures. *Neurotox. Res.* 18, 377–385.

Bell C, Dinwiddie D, Miller N, Hateley S, Ganusova E, Mudge J, Langley R, Zhang L, Lee C, Schilkey F, Sheth V, Woodward J, Peckham H, Schroth G, Kim R *et al.* (2011) Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Sci. Transl. Med.* 3, 65ra4.

Benedix A, Becker C, de Groot B, Caflisch A, and Böckmann R (2009) Predicting free energy changes using structural ensembles. *Nat. Methods* 6, 3–4.

Berkel S, Marshall C, Weiss B, Howe J, Roeth R, Moog U, Endris V, Roberts W, Szatmari P, Pinto D, Bonin M, Riess A, Engels H, Sprengel R, Scherer S *et al.* (2010) Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nat. Genet.* 42, 489–491.

Berry-Kravis E, Hessl D, Coffey S, Hervey C, Schneider A, Yuhas J, Hutchison J, Snape M, Tranfaglia M, Nguyen DV, and Hagerman R (2009) A pilot open label, single dose trial of fenobam in adults with fragile X syndrome. *J. Med. Genet.* 46, 266–271.

Berry-Kravis E, Sumis A, Hervey C, Nelson M, Porges SW, Weng N, Weiler IJ, and Greenough WT (2008) Open-label treatment trial of lithium to target the underlying defect in fragile X syndrome. *J. Dev. Behav. Pediatr.* 29, 293–302.

Bhakar AL, Dölen G, and Bear MF (2012) The Pathophysiology of Fragile X (and What It Teaches Us about Synapses). *Annu. Rev. Neurosci.* 35, 417–443.

Bhalla K, Luo Y, Buchan T, Beachem MA, Guzauskas GF, Ladd S, Bratcher SJ, Schroer RJ, Balsamo J, DuPont BR, Lilien J, and Srivastava AK (2008) Alterations in CDH15 and KIRREL3 in patients with mild to severe intellectual disability. *Am. J. Hum. Genet.* 83, 703–713.

Biedler J, Roffler-Tarlov S, Schachner M, and Freedman L (1978) Multiple neurotransmitter synthesis by human neuroblastoma cell lines and clones. *Cancer Res.* 38, 3751–3757.

Biou V, Yaremchuk A, Tukalo M, and Cusack S (1994) The 2.9 A crystal structure of T. thermophilus seryl-tRNA synthetase complexed with tRNA(Ser). *Science* 263, 1404–1410.

Blaner W and Churchich J (1979) Succinic semialdehyde dehydrogenase. Reactivity of lysyl residues. *J. Biol. Chem.* 254, 1794–1798.

Borel F, Vincent C, Leberman R, and Härtlein M (1994) Seryl-tRNA synthetase from Escherichia coli: implication of its N-terminal domain in aminoacylation activity and specificity. *Nucleic Acids Res.* 22, 2963–2969.

Bull L, van Eijk MJT, Pawlikowska L, DeYoung J, Juijn J, Liao M, Klomp LWJ, Lomri N, Berger R, Scharschmidt B, Knisely A, Houwen RHJ, and Freimer N (1998) A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. *Nat. Genet.* 18, 219–224.

Buzzi A, Wu Y, Frantseva MV, Perez Velazquez JL, Cortez MA, Liu CC, Shen LQ, Gibson KM, and Snead O (2006) Succinic semialdehyde dehydrogenase deficiency: GABAB receptor-mediated function. *Brain Res.* 1090, 15–22.

Cabeza R and Nyberg L (2000) Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12, 1–47.

Caito S, Milatovic D, Hill K, Aschner M, Burk R, and Valentine W (2011) Progression of neurodegeneration and morphologic changes in the brains of juvenile mice with selenoprotein P deleted. *Brain Res.* 1398, 1–12.

Capriotti E, Fariselli P, Calabrese R, and Casadio R (2005a) Predicting protein stability changes from sequences using support vector machines. *Bioinformatics* 21, ii54–58.

Capriotti E, Fariselli P, and Casadio R (2005b) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 33, W306–310.

Carlisle HJ, Manzerra P, Marcora E, and Kennedy MB (2008) SynGAP regulates steady-state and activity-dependent phosphorylation of cofilin. *J. Neurosci.* 28, 13 673–13 683.

Carter JCW (1993) Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.* 62, 715–748.

Casey JP, McGettigan P, Lynam-Lennon N, McDermott M, Regan R, Conroy J, Bourke B, Sullivan JO, Crushell E, Lynch S, and Ennis S (2012) Identification of a mutation in LARS as a novel cause of infantile hepatopathy. *Mol. Genet. Metab.* 106, 351–358.

Cash C, Ciesielski L, Maitre M, and Mandel P (1977) Purification and properties of rat brain succinic semialdehyde dehydrogenase. *Biochimie* 59, 257–268.

Chelly J and Mandel JL (2001) Monogenic causes of X-linked mental retardation. *Nat. Rev. Genet.* 2, 669–680.

Cheng J, Randall A, and Baldi P (2006) Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins.* 62, 1125–1132.

Cho G, Nasrallah MP, Lim Y, and Golden JA (2012) Distinct DNA binding and transcriptional repression characteristics related to different ARX mutations. *Neurogenetics* 13, 23–29.

Choi CH, Schoenfeld BP, Bell AJ, Hinchey P, Kollaros M, Gertner MJ, Woo NH, Tranfaglia MR, Bear MF, Zukin RS, McDonald TV, Jongens TA, and McBride SMJ (2011) Pharmacological reversal of synaptic plasticity deficits in the mouse model of fragile X syndrome by group II mGluR antagonist or lithium treatment. *Brain Res.* 1380, 106–119.

Choi Y, Sims G, Murphy S, Miller J, and Chan A (2012) Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 7, e46 688.

Clarke H, Dalley J, Crofts H, Robbins T, and Roberts A (2004) Cognitive inflexibility after prefrontal serotonin depletion. *Science* 304, 878–880.

Collins SC, Bray SM, Suhl JA, Cutler DJ, Coffee B, Zwick ME, and Warren ST (2010) Identification of novel FMR1 variants by massively parallel sequencing in developmentally delayed males. *Am. J. Med. Genet. A.* 152A, 2512–2520.

Comery TA, Harris JB, Willems PJ, Oostra BA, Irwin SA, Weiler IJ, and Greenough WT (1997) Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits. *Proc. Natl. Acad. Sci. USA.* 94, 5401–5404.

Coordinators NR (2013) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 41, D8–D20.

Corbetta M and Shulman G (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Costa-Mattioli M, Sonenberg N, and Richter J (2009) Translational regulatory mechanisms in synaptic plasticity and memory storage. *Prog. Mol. Biol. Transl. Sci.* 90, 293–311.

Cusack S (1997) Aminoacyl-tRNA synthetases. *Curr. Opin. Struct. Biol.* 7, 881–889.

Cusack S (1999) RNA-protein complexes. *Curr. Opin. Struct. Biol.* 9, 66–73.

De Boulle K, Verkerk A, Reyniers E, Vits L, Hendrickx J, Van Roy B, Van den Bos F, de Graaff E, Oostra B, and Willems P (1993) A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat. Genet.* 3, 31–35.

Dehouck Y, Kwasigroch J, Gilis D, and Rooman M (2005) PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics* 12, 151.

Doherty J, Hattox S, Snead O, and Roth R (1978) Identification of endogenous gamma-hydroxybutyrate in human and bovine brain and its regional distribution in human, guinea pig and rhesus monkey brain. *J. Pharmacol. Exp. Ther.* 207, 130–139.

Dölen G, Osterweil E, Rao B, Smith G, Auerbach B, Chattarji S, and Bear M (2007) Correction of fragile X syndrome in mice. *Neuron* 56, 955–962.

Dosztanyi Z, Magyar C, Tusnady G, and Simon I (2003) SCide: identification of stabilization centers in proteins. *Bioinformatics.* 19, 899–900.

Duncan R and Tipton K (1971) The kinetics of pig brain aldehyde dehydrogenase. *Eur. J. Biochem.* 22, 538–543.

Estrada-Bernal A, Sanford SD, Sosa LJ, Simon GC, Hansen KC, and Pfenninger KH (2012) Functional complexity of the axonal growth cone: a proteomic analysis. *PloS One* 7, e31 858.

Farlow MR (2009) Treatment of mild cognitive impairment (MCI). *Curr. Alzheimer Res.* 6, 362–367.

Feuk L, Carson A, and Scherer S (2006) Structural variation in the human genome. *Nat. Rev. Genet.* 7, 85–97.

Finn R, Mistry J, Tate J, Coggill P, Heger A, Pollington J, Gavin O, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer E, Eddy S, and Bateman A (2010) The Pfam protein families database. *Nucleic Acids Res.* 38, D211–222.

Fredman D, White SJ, Potter S, Eichler EE, Den Dunnen JT, and Brookes AJ (2004) Complex SNP-related sequence variation in segmental genome duplications. *Nat. Genet.* 36, 861–866.

Freedman D, Riesenhuber M, Poggio T, and Miller E (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291, 312–316.

Friocourt G and Parnavelas JG (2011) Identification of Arx targets unveils new candidates for controlling cortical interneuron migration and differentiation. *Front. Cell Neurosci.* 5, 28.

Froestl W, Gallagher M, Jenkins H, Madrid A, Melcher T, Teichman S, Mondadori CG, and Pearlman R (2004) SGS742: the first GABA(B) receptor antagonist in clinical trials. *Biochem. Pharmacol.* 68, 1479–1487.

Frye M and Watt FM (2006) The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr. Biol.* 16, 971–981.

Fukui H, Hanaoka R, and Kawahara A (2009) Noncanonical activity of seryl-tRNA synthetase is involved in vascular development. *Circ. Res.* 104, 1253–1259.

Gal-Ben-Ari S, Kenney J, Ounalla-Saad H, E T, David O, Levitan D, Gildish I, Panja D, Pa 1B, Wibrand K, Simpson T, Proud C, Bramham C, Armstrong J, and Rosenblum K (2012) Consolidation and translation regulation. *Learn Mem.* 19, 410–422.

Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, and Mackay JP (2007) Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem. Sci.* 32, 63–70.

Garshasbi M, Hadavi V, Habibi H, Kahrizi K, Kariminejad R, Behjati F, Tzschach A, Najmabadi H, Ropers H, and Kuss A (2008) A defect in the TUSC3 gene is associated with autosomal recessive mental retardation. *Am. J. Hum. Genet.* 82, 1158–1164.

Garshasbi M, Kahrizi K, Hosseini M, Nouri Vahid L, Falah M, Hemmati S, Hu H, Tzschach A, Ropers HH, Najmabadi H, and Kuss AW (2011) A novel nonsense mutation in TUSC3 is responsible for non-syndromic autosomal recessive mental retardation in a consanguineous Iranian family. *Am. J. Med. Genet. A.* 155A, 1976–1980.

Garshasbi M, Motazacker MM, Kahrizi K, Behjati F, Abedini SS, Nieh SE, Firouzabadi SG, Becker C, Rüschendorf F, Nürnberg P, Tzschach A, Vazifehmand R, Erdogan F, Ullmann R, Lenzner S *et al.* (2006) SNP array-based homozygosity mapping reveals MCPH1 deletion in family with autosomal recessive mental retardation and mild microcephaly. *Hum. Genet.* 118, 708–715.

Gécz J, Cloosterman D, and Partington M (2006) ARX: a gene for all seasons. *Curr. Opin. Genet. Dev.* 16, 308–316.

Gécz J, Shoubridge C, and Corbett M (2009) The genetic landscape of intellectual disability arising from chromosome X. *Trends Genet.* 25, 308–316.

Gervasi N, Monnier Z, Vincent P, Paupardin-Tritsch D, Hughes SW, Crunelli V, and Leresche N (2003) Pathway-specific action of gamma-hydroxybutyric acid in sensory thalamus and its relevance to absence seizures. *J. Neurosci.* 23, 11 469–11 478.

Gey G, Coffman W, and Kubicek M (1952) Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. *Cancer Res.* 12, 264–265.

Gibson K, Aramaki S, Sweetman L, Nyhan W, DeVivo D, Hodson A, and Jakobs C (1990) Stable isotope dilution analysis of 4-hydroxybutyric acid: an accurate method for quantification in physiological fluids and the prenatal diagnosis of 4-hydroxybutyric aciduria. *Biomed. Environ. Mass Spectrom.* 19, 89–93.

Gibson K, Christensen E, Jakobs C, Fowler B, Clarke M, Hammersen G, Raab K, Kobori J, Moosa A, Vollmer B, Rossier E, Iafolla A, Matern D, Brouwer O, Finkelstein J *et al.* (1997) The clinical phenotype of succinic semialdehyde dehydrogenase deficiency (4-hydroxybutyric aciduria): case reports of 23 new patients. *Pediatrics* 99, 567–574.

Gibson K, Lee C, Chambliss K, Kamali V, Francois B, Jaeken J, and Jakobs C (1991) 4-Hydroxybutyric aciduria: application of a fluorometric assay to the determination of succinic semialdehyde dehydrogenase activity in extracts of cultured human lymphoblasts. *Clin. Chim. Acta.* 196, 219–221.

Gibson KM, Gupta M, Pearl PL, Tuchman M, Vezina LG, Snead rOC, Smit LME, and Jakobs C (2003) Significant behavioral disturbances in succinic semialdehyde dehydrogenase (SSADH) deficiency (gamma-hydroxybutyric aciduria). *Biol. Psychiatry* 54, 763–768.

Graham F, Smiley J, Russell W, and Nairn R (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36, 59–74.

Griggs BL, Ladd S, Saul RA, DuPont BR, and Srivastava AK (2008) Dedicator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. *Genomics* 91, 195–202.

Gropman A (2003) Vigabatrin and newer interventions in succinic semialdehyde dehydrogenase deficiency. *Ann. Neurol.* 54 Suppl 6, S66–72.

Grosshans H, Hurt E, and Simos G (2000) An aminoacylation-dependent nuclear tRNA export pathway in yeast. *Genes Dev.* 14, 830–840.

Gruber A, Calhoon G, Shusterman I, Schoenbaum G, Roesch M, and O'Donnell P (2010) More is less: a disinhibited prefrontal cortex impairs cognitive flexibility. *J. Neurosci.* 30, 17 102–17 110.

Guo X, Hamilton PJ, Reish NJ, Sweatt JD, Miller CA, and Rumbaugh G (2009) Reduced expression of the NMDA receptor-interacting protein SynGAP causes behavioral abnormalities that model symptoms of Schizophrenia. *Neuropsychopharmacology* 34, 1659–1672.

Gupta M, Greven R, Jansen EEW, Jakobs C, Hogema BM, Froestl W, Snead OC, Bartels H, Grompe M, and Gibson KM (2002) Therapeutic intervention in mice deficient for succinate semialdehyde dehydrogenase (gamma-hydroxybutyric aciduria). *J. Pharmacol. Exp. Ther.* 302, 180–187.

Hagerman R, Lauterborn J, Au J, and Berry-Kravis E (2012) Fragile X syndrome and targeted treatment trials. *Results. Probl. Cell. Differ.* 54, 297–335.

Hamdan FF, Gauthier J, Araki Y, Lin DT, Yoshizawa Y, Higashi K, Park AR, Spiegelman D, Dobrzeniecka S, Piton A, Tomitori H, Daoud H, Massicotte C, Henrion E, Diallo O *et al.* (2011a) Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* 88, 306–316.

Hamdan FF, Gauthier J, Dobrzeniecka S, Lortie A, Mottron L, Vanasse M, D'Anjou G, Lacaille JC, Rouleau GA, and Michaud JL (2011b) Intellectual disability without epilepsy associated with STXBP1 disruption. *Eur. J. Hum. Genet.* 19, 607–609.

Hamdan FF, Gauthier J, Spiegelman D, Noreau A, Yang Y, Pellerin S, Dobrzeniecka S, Côté M, Perreau-Linck E, Perreault-Linck E, Carmant L, D'Anjou G, Fombonne E, Addington AM, Rapoport JL *et al.* (2009) Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. *N. Engl. J. Med.* 360, 599–605.

Herzog W, Müller K, Huisken J, and Stainier D (2009) Genetic evidence for a noncanonical function of seryl-tRNA synthetase in vascular development. *Circ. Res.* 104, 1260–1266.

van den Heuvel M, Stam C, Kahn R, and Hulshoff Pol H (2009) Efficiency of functional brain networks and intellectual performance. *J. Neurosci.* 29, 7619–7624.

Hinton VJ, Brown WT, Wisniewski K, and Rudelli RD (1991) Analysis of neocortex in three males with the fragile X syndrome. *Am. J. Med. Genet.* 41, 289–294.

Hochstenbach R, van Binsbergen E, Engelen J, Nieuwint A, Polstra A, Poddighe P, Ruivenkamp C, Sikkema-Raddatz B, Smeets D, and Poot M (2009) Array analysis and karyotyping: workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherlands. *Eur. J. Med. Genet.* 52, 161–169.

Hogema B, M G, Senephansiri H, Burlingame T, Taylor M, Jakobs C, Schutgens R, Froestl W, Snead O, Diaz-Arrastia R, Bottiglieri T, Grompe M, and Gibson K (2001) Pharmacologic rescue of lethal seizures in mice deficient in succinate semialdehyde dehydrogenase. *Nat. Genet.* 29, 212–216.

Hoyer J, Ekici AB, Endele S, Popp B, Zweier C, Wiesener A, Wohlleber E, Dufke A, Rossier E, Petsch C, Zweier M, Göhring I, Zink AM, Rappold G, Schröck E *et al.* (2012) Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am. J. Hum. Genet.* 90, 565–572.

Hu H, Eggers K, Chen W, Garshasbi M, Motazacker MM, Wrogemann K, Kahrizi K, Tzschach A, Hosseini M, Bahman I, Hucho T, Mühlenhoff M, Gerardy-Schahn R, Najmabadi H, Ropers HH *et al.* (2011) ST3GAL3 Mutations Impair the Development of Higher Cognitive Functions. *Am. J. Hum. Genet.* 89, 407–414.

Hu H, Wrogemann K, Kalscheuer V, Tzschach A, Richard H, Haas S, Menzel C, Bienek M, Froyen G, Raynaud M, Van Bokhoven H, Chelly J, Ropers H, and Chen W (2009) Mutation screening in 86 known X-linked mental retardation genes by droplet-based multiplex PCR and massive parallel sequencing. *Hugo J.* 3, 41–49.

Hu WH, Pendergast JS, Mo XM, Brambilla R, Bracchi-Ricard V, Li F, Walters WM, Blits B, He L, Schaal SM, and Bethea JR (2005) NIBP, a novel NIK and IKK(beta)-binding protein that enhances NF-(kappa)B activation. *J. Biol. Chem.* 280, 29 233–29 241.

Hussain S, Benavente S, Nascimento E, Dragoni I, Kurowski A, Gillich A, Humphreys P, and Frye M (2009) The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. *J. Cell Biol.* 186, 27–40.

Ibba M, Curnow AW, and Söll D (1997) Aminoacyl-tRNA synthesis: divergent routes to a common goal. *Trends Biochem. Sci.* 22, 39–42.

Ibba M and Soell D (2000) Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* 69, 617–650.

Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee Y, Narzisi G, Leotta A, Kendall J, Grabowska E, Ma B, Marks S, Rodgers L *et al.* (2012) De novo gene disruptions in children on the autistic spectrum. *Neuron* 74, 285–299.

Irwin SA, Galvez R, and Greenough WT (2000) Dendritic spine structural anomalies in fragile-X mental retardation syndrome. *Cereb. Cortex* 10, 1038–1044.

Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P, and Gasteiger E (2009) Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinformatics* 10, 136.

Jakobs C, Jaeken J, and Gibson K (1993) Inherited disorders of GABA metabolism. *J. Inherit. Metab. Dis.* 16, 704–715.

Jansen EEW, Struys E, Jakobs C, Hager E, Snead OC, and Gibson KM (2008) Neurotransmitter alterations in embryonic succinate semialdehyde dehydrogenase (SSADH) deficiency suggest a heightened excitatory state during development. *BMC Dev. Biol.* 8, 112.

Jia J, Arif A, Ray P, and Fox P (2008) WHEP domains direct noncanonical function of glutamyl-prolyl tRNA synthetase in translational control of gene expression. *Mol. Cell.* 29, 679–690.

Jordanova A, Irobi J, Thomas F, Van Dijck P, Meerschaert K, Dewil M, Dierick I, Jacobs A, De Vriendt E, Guergueltcheva V, Rao C, Tournev I, Gondim F, D'Hooghe M, Van Gerwen V *et al.* (2006) Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. *Nat. Genet.* 38, 197–202.

Jung H and Holt C (2011) Local translation of mRNAs in neural development. *Wiley Interdiscip. Rev. RNA.* 2, 153–165.

Kammeraat C and Veldstra H (1968) Characterization of succinate semialdehyde dehydrogenase from rat brain. *Biochim. Biophys. Acta.* 151, 1–10.

Kang J, Park Y, Huh T, Lee W, Choi M, and Kwon O (2005) High-level expression and characterization of the recombinant enzyme, and tissue distribution of human succinic semialdehyde dehydrogenase. *Protein. Expr. Purif.* 44, 16–22.

Karaveg K, Siriwardena A, Tempel W, Liu Z, Glushka J, Wang B, and Moremen K (2005) Mechanism of class 1 (glycosylhydrolase family 47) alpha-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control. *J. Biol. Chem.* 280, 16 197–16 207.

Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, and Kent WJ (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* 32, D493–496.

Kaufman L, Ayub M, and Vincent JB (2010) The genetic basis of non-syndromic intellectual disability: a review. *J. Neurodev. Disord.* 2, 182–209.

Kelly V, Sherratt P, Crouch D, and Hayes J (2002) Novel homodimeric and heterodimeric rat gamma-hydroxybutyrate synthases that associate with the Golgi apparatus define a distinct subclass of aldo-keto reductase 7 family proteins. *Biochem. J.* 366, 847–861.

Kennedy GC, Matsuzaki H, Dong S, Liu Wm, Huang J, Liu G, Su X, Cao M, Chen W, Zhang J, Liu W, Yang G, Di X, Ryder T, He Z *et al.* (2003) Large-scale genotyping of complex DNA. *Nat. Biotechnol.* 21, 1233–1237.

Kent W, Sugnet C, Furey T, Roskin K, Pringle T, Zahler A, and Haussler D (2002) The Human Genome Browser at UCSC. *Genome Res.* 12, 996–1006.

Khan MA, Rafiq MA, Noor A, Ali N, Ali G, Vincent JB, and Ansar M (2011) A novel deletion mutation in the TUSC3 gene in a consanguineous Pakistani family with autosomal recessive nonsyndromic intellectual disability. *BMC Med. Genet.* 12, 56.

Kim J, Lee HK, Takamiya K, and Huganir R (2003) The Role of Synaptic GTPase-Activating Protein in Neuronal Development and Synaptic Plasticity. *J. Neurosci.* 23, 1119–1124.

Kim KJ, Pearl P, Jensen K, Snead OC, Malaspina P, Jakobs C, and Gibson KM (2011) Succinic Semialdehyde Dehydrogenase: Biochemical-Molecular-Clinical Disease Mechanisms, Redox Regulation, and Functional Significance. *Antioxid. Redox. Signal.* 15, 691–718.

Kim YG, Lee S, Kwon OS, Park SY, Lee SJ, Park BJ, and Kim KJ (2009) Redox-switch modulation of human SSADH by dynamic catalytic loop. *EMBO J.* 28, 959–968.

Kingsmore S (2012) Comprehensive carrier screening and molecular diagnostic testing for recessive childhood diseases. *PLoS Curr.* 2, e4f9877ab8ffa9.

Kingsmore S, Dinwiddie D, Miller N, Soden S, and Saunders C (2011) Adopting orphans: comprehensive genetic testing of Mendelian diseases of childhood by next-generation sequencing. *Expert Rev. Mol. Diagn.* 11, 855–868.

Knerr I, Pearl P, Bottiglieri T, Snead O, Jakobs C, and Gibson K (2007) Therapeutic concepts in succinate semialdehyde dehydrogenase (SSADH; ALDH5a1) deficiency (gamma-hydroxybutyric aciduria). Hypotheses evolved from 25 years of patient evaluation, studies in Aldh5a1-/- mice and characterization of gamma-hydroxybutyric acid pharmacology. *J. Inherit. Metab. Dis.* 30, 279–294.

Knuesel I, Elliott A, Chen HJ, Mansuy IM, and Kennedy MB (2005) A role for synGAP in regulating neuronal apoptosis. *Eur. J. Neurosci.* 21, 611–621.

Komiyama N, Watabe A, Carlisle H, Porter K, Charlesworth P, Monti J, Strathdee D, O'Carroll C, Martin S, Morris R, O'Dell T, and Grant S (2002) SynGAP Regulates ERK/MAPK Signaling, Synaptic Plasticity, and Learning in the Complex with Postsynaptic Density 95 and NMDA Receptor. *J. Neurosci.* 22.

Konovalova S and Tyynismaa H (2013) Mitochondrial aminoacyl-tRNA synthetases in human disease. *Mol. Genet. Metab.* pii: S1096-7192, 00 032–2.

Kramer J and van Bokhoven H (2009) Genetic and epigenetic defects in mental retardation. *Int. J. Biochem. Cell Biol.* 41, 96–107.

Krueger DD, Osterweil EK, Chen SP, Tye LD, and Bear MF (2011) Cognitive dysfunction and prefrontal synaptic abnormalities in a mouse model of fragile X syndrome. *Proc. Natl. Acad. Sci. USA.* 108, 2587–2592.

Kumar P, Henikoff S, and Ng P (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4, 1073–1081.

Lander E and Botstein D (1987) Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 236, 1567–1570.

Laumonnier F, Cuthbert PC, and Grant SGN (2007) The role of neuronal complexes in human X-linked brain diseases. *Am. J. Hum. Genet.* 80, 205–220.

Leonard H and Wen X (2002) The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment. Retard. Dev. Disabil. Res. Rev.* 8, 117–134.

Li Y, Vinckenbosch N, Tian G, Huerta-Sanchez E, Jiang T, Jiang H, Albrechtsen A, Andersen G, Cao H, Korneliussen T, Grarup N, Guo Y, Hellman I, Jin X, Li Q *et al.* (2010) Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. *Nat. Genet.* advance online publication, 969–972.

de Ligt J, Willemsen MH, van Bon BWM, Kleefstra T, Yntema HG, Kroes T, Vultovan Silfhout AT, Koolen DA, de Vries P, Gilissen C, Del Rosario M, Hoischen A, Scheffer H, de Vries BBA, Brunner HG *et al.* (2012) Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N. Engl. J. Med.* 367, 1921–1929.

Linden D (2007) The working memory networks of the human brain. *Neuroscientist.* 13, 257–267.

Lisik M (2010) Genetics of X-linked Mental Retardation. *Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester* .

Liu ZH, Huang T, and Smith CB (2012) Lithium reverses increased rates of cerebral protein synthesis in a mouse model of fragile X syndrome. *Neurobiol. Dis.* 45, 1145–1152.

Loane M, Morris JK, Addor MC, Arriola L, Budd J, Doray B, Garne E, Gatt M, Haeusler M, Khoshnood B, Klungsøyr Melve K, Latos-Bielenska A, McDonnell B, Mullaney C, O'Mahony M *et al.* (2012) Twenty-year trends in the prevalence of Down syndrome and other trisomies in Europe: impact of maternal age and prenatal screening. *Eur. J. Hum. Genet.* 21, 27–33.

Loerch P, Lu T, Dakin K, Vann J, Isaacs A, Geula C, Wang J, Pan Y, Gabuzda D, Li C, Prolla T, and Yankner B (2008) Evolution of the aging brain transcriptome and synaptic regulation. *PLoS One.* 3, e3329.

Lu J and Holmgren A (2009) Selenoproteins. *J. Biol. Chem.* 284, 723–727.

Lu T, Pan Y, Kao SY, Li C, Kohane I, Chan J, and Yankner B (2004) Gene regulation and DNA damage in the ageing human brain. *Nature* 429, 883–891.

Lubs HA, Stevenson RE, and Schwartz CE (2012) Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* 90, 579–590.

Lugenbeel K, Peier A, Carson N, Chudley A, and Nelson D (1995) Intragenic loss of function mutations demonstrate the primary role of FMR1 in fragile X syndrome. *Nat. Genet.* 10, 483–485.

Lund E and Dahlberg JE (1998) Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science* 282, 2082–2085.

Lupski J, Reid J, Gonzaga-Jauregui C, Rio Deiros D, Chen D, Nazareth L, Bainbridge M, Dinh H, Jing C, Wheeler D, McGuire A, Zhang F, Stankiewicz P, Halperin J, Yang C *et al.* (2010) Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N. Engl. J. Med.* 362, 1181–1191.

Lyon R, Johnston S, Watson D, McGarvie G, and Ellis E (2007) Synthesis and catabolism of gamma-hydroxybutyrate in SH-SY5Y human neuroblastoma cells: role of the aldo-keto reductase AKR7A2. *J. Biol. Chem.* 282, 25 986–25 992.

Maitre M (1997) The gamma-hydroxybutyrate signalling system in brain: organization and functional implications. *Prog. Neurobiol.* 51, 337–361.

Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, and Jabado N (2011) What can exome sequencing do for you? *J. Med. Genet.* 48, 580–589.

Marangi G, Leuzzi V, Manti F, Lattante S, Orteschi D, Pecile V, Neri G, and Zollino M (2012) TRAPPC9-related autosomal recessive intellectual disability: report of a new mutation and clinical phenotype. *Eur. J. Hum. Genet.* 21, 229–232.

Martinez FJ, Lee JH, Lee JE, Blanco S, Nickerson E, Gabriel S, Frye M, Al-Gazali L, and Gleeson JG (2012) Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *J. Med. Genet.* 49, 380–385.

Matsuzaki H, Loi H, Dong S, Tsai YY, Fang J, Law J, Di X, Liu WM, Yang G, Liu G, Huang J, Kennedy GC, Ryder TB, Marcus GA, Walsh PS *et al.* (2004) Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res.* 14, 414–425.

McLaren J and Bryson SE (1987) Review of recent epidemiological studies of mental retardation: prevalence, associated disorders, and etiology. *Am. J. Ment. Retard.* 92, 243–254.

McLaughlin HM, Sakaguchi R, Liu C, Igarashi T, Pehlivan D, Chu K, Iyer R, Cruz P, Cherukuri PF, Hansen NF, Mullikin JC, Biesecker LG, Wilson TE, Ionasescu V, Nicholson G *et al.* (2010) Compound heterozygosity for loss-of-function lysyl-tRNA synthetase mutations in a patient with peripheral neuropathy. *Am. J. Hum. Genet.* 87, 560–566.

Miller J, McLachlan AD, and Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *EMBO J.* 4, 1609–1614.

Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, Kahrizi K, Rafiq MA, Weksberg R, Nasr T, Naeem F, Tzschach A, Kuss AW, Ishak GE, Doherty D *et al.* (2009) Identification of mutations in TRAPPC9, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation. *Am. J. Hum. Genet.* 85, 909–915.

Mochida GH, Mahajnah M, Hill AD, Basel-Vanagaite L, Gleason D, Hill RS, Bodell A, Crosier M, Straussberg R, and Walsh CA (2009) A truncating mutation of TRAPPC9 is associated with autosomal-recessive intellectual disability and postnatal microcephaly. *Am. J. Hum. Genet.* 85, 897–902.

Modell B and Darr A (2002) Science and society: genetic counselling and customary consanguineous marriage. *Nat. Rev. Genet.* 3, 225–229.

Molinari F, Foulquier F, Tarpey PS, Morelle W, Boissel S, Teague J, Edkins S, Futreal PA, Stratton MR, Turner G, Matthijs G, Gecz J, Munnich A, and Colleaux L (2008) Oligosaccharyltransferase-subunit mutations in nonsyndromic mental retardation. *Am. J. Hum. Genet.* 82, 1150–1157.

Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini S, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, Püttmann L, Vahid L, Jensen C *et al.* (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 478, 57–63.

Ng P and Henikoff S (2006) Predicting the effects of amino acid substitutions on protein function. *Annu. Rev. Genomics Hum. Genet.* 7, 61–80.

Ng PC and Henikoff S (2001) Predicting deleterious amino acid substitutions. *Genome Res.* 11, 863–874.

Ng PC and Henikoff S (2002) Accounting for human polymorphisms predicted to affect protein function. *Genome Res.* 12, 436–446.

Ng PC and Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31, 3812–3814.

O'Donnell M, Chance RK, and Bashaw GJ (2009) Axon growth and guidance: receptor regulation and signal transduction. *Annu. Rev. Neurosci.* 32, 383–412.

O'Roak B, Vives L, Girirajan S, Karakoc E, Krumm N, Coe B, Levy R, Ko A, Lee C, Smith J, Turner E, Stanaway I, Vernot B, Malig M, Baker C *et al.* (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485, 246–250.

Pabo CO, Peisach E, and Grant RA (2001) Design and selection of novel Cys2His2 zinc finger proteins. *Annu. Rev. Biochem.* 70, 313–340.

Pasupathy A and Miller E (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876.

Pearl P, Gibson K, Acosta M, Vezina L, Theodore W, Rogawski M, Novotny E, Gropman A, Conry J, Berry G, and Tuchman M (2003) Clinical spectrum of succinic semialdehyde dehydrogenase deficiency. *Neurology* 60, 1413–1417.

Pearl P, Shukla L, Theodore W, Jakobs C, and KM G (2011) Epilepsy in succinic semialdehyde dehydrogenase deficiency, a disorder of GABA metabolism. *Brain Dev.* 33, 796–805.

Pearl PL, Gibson KM, Quezado Z, Dustin I, Taylor J, Trzcinski S, Schreiber J, Forester K, Reeves-Tyer P, Liew C, Shamim S, Herscovitch P, Carson R, Butman J, Jakobs C *et al.* (2009) Decreased GABA-A binding on FMZ-PET in succinic semialdehyde dehydrogenase deficiency. *Neurology* 73, 423–429.

Pfundt R and Veltman JA (2012) Structural genomic variation in intellectual disability. *Methods Mol. Biol.* 838, 77–95.

Philippe O, Rio M, Carioux A, Plaza JM, Guigue P, Molinari F, Boddaert N, Bole-Feysot C, Nitschke P, Smahi A, Munnich A, and Colleaux L (2009) Combination of linkage mapping and microarray-expression analysis identifies NF-kappaB signaling defect as a cause of autosomal-recessive mental retardation. *Am. J. Hum. Genet.* 85, 903–908.

Ponten J and Macintyre E (1968) Long term culture of normal and neoplastic human glia. *Acta Pathol. Microbiol. Scand.* 74, 465–86.

Porter R, Jaeschke G, Spooren W, Ballard TM, Büttelmann B, Kolczewski S, Peters JU, Prinssen E, Wichmann J, Vieira E, Mühlemann A, Gatti S, Mutel V, and Malherbe P (2005) Fenobam: a clinically validated nonbenzodiazepine anxiolytic is a potent, selective, and noncompetitive mGlu5 receptor antagonist with inverse agonist activity. *J. Pharmacol. Exp. Ther.* 315, 711–721.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, and Lopez R (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* 33, W116–120.

Rafiq MA, Kuss AW, Puettmann L, Noor A, Ramiah A, Ali G, Hu H, Kerio NA, Xiang Y, Garshasbi M, Khan MA, Ishak GE, Weksberg R, Ullmann R, Tzschach A *et al.* (2011) Mutations in the alpha 1,2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* 89, 176–182.

Rapaport E, Remy P, Kleinkauf H, Vater J, and Zamecnik P (1987) Aminoacyl-tRNA synthetases catalyze AMP – ADP – ATP exchange reactions, indicating labile covalent enzyme-amino-acid intermediates. *Proc. Natl. Acad. Sci. USA.* 84, 7891–7895.

Rauch A, Hoyer J, Guth S, Zweier C, Kraus C, Becker C, Zenker M, Hüffmeier U, Thiel C, Rüschendorf F, Nürnberg P, Reis A, and Trautmann U (2006) Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am. J. Med. Genet. A.* 140, 2063–2074.

Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, Dufke A, Cremer K, Hempel M, Horn D, Hoyer J *et al.* (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 380, 1674–1682.

Ray PS and Fox PL (2007) A post-transcriptional pathway represses monocyte VEGF-A expression and angiogenic activity. *EMBO J.* 26, 3360–3372.

Raymond FL and Tarpey P (2006) The genetics of mental retardation. *Hum. Mol. Genet.* 15, R110–6.

Redon R, Ishikawa S, Fitch K, Feuk L, Perry G, Andrews TD, Fiegler H, Shapero M, Carson A, Chen W, Cho EK, Dallaire S, Freeman J, Gonzalez J, Gratacos M *et al.* (2006) Global variation in copy number in the human genome. *Nature* 444, 444–454.

Rivett A and Tipton K (1981) Kinetic studies with rat-brain succinic-semialdehyde dehydrogenase. *Eur. J. Biochem.* 117, 187–193.

Robertson L (2002) Memory and the brain. *J. Dent. Educ.* 66, 30–42.

Ropers H (2006) X-linked mental retardation: many genes for a complex disorder. *Curr. Opin. Genet. Dev.* 16, 260–269.

Ropers H and Hamel B (2005) X-linked mental retardation. *Nat. Rev. Genet.* 6, 46–57.

Ropers HH (2008) Genetics of intellectual disability. *Curr. Opin. Genet. Dev.* 18, 241–250.

Ropers HH (2010a) Genetics of early onset cognitive impairment. *Annu. Rev. Genomics. Hum. Genet.* 11, 161–187.

Ropers HH (2010b) Single gene disorders come into focus–again. *Dialogues Clin. Neurosci.* 12, 95–102.

Ropers HH (2012) On the future of genetic risk assessment. *J. Community. Genet.* 3, 229–236.

Ross M, Grafham D, Coffey A, Scherer S, McLay K, Muzny D, Platzer M, Howell G, Burrows C, Bird C, Frankish A, Lovell F, Howe K, Ashurst J, Fulton R *et al.* (2005) The DNA sequence of the human X chromosome. *Nature* 434, 325–337.

Rozen S and Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386.

Rygula R, Walker S, Clarke H, Robbins T, and Roberts A (2010) Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *J. Neurosci.* 30, 14 552–14 559.

Sacher M, Kim YG, Lavie A, Oh BH, and Segev N (2008) The TRAPP complex: insights into its architecture and function. *Traffic* 9, 2032–2042.

Sanders S, Murtha M, Gupta A, Murdoch J, Raubeson M, Willsey A, Ercan-Sencicek A, DiLullo N, Parikshak N, Stein J, Walker M, Ober G, Teran N, Song Y, El-Fishawy P *et al.* (2012) De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485, 237–241.

Santen GWE, Aten E, Sun Y, Almomani R, Gilissen C, Nielsen M, Kant SG, Snoeck IN, Peeters EAJ, Hilhorst-Hofstee Y, Wessels MW, den Hollander NS, Ruivenkamp CAL, van Ommen GJB, Breuning MH *et al.* (2012) Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* 44, 379–380.

Sarkar S, Azad A, and Hopper A (1999) Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* 96, 14 366–14 371.

Saunders C, Miller N, Soden S, Dinwiddie D, Noll A, Alnadi N, Andraws N, Patterson M, Krivohlavek L, Fellis J, Humphray S, Saffrey P, Kingsbury Z, Weir J, Betley J *et al.* (2012) Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Sci. Transl. Med.* 4, 154ra135.

Scherer W and Hoogasian A (1954) Preservation at subzero temperatures of mouse fibroblasts (strain L) and human epithelial cells (strain HeLa). *Proc. Soc. Exp. Biol. Med.* 87, 480–487.

Schimmel P (1991) Classes of aminoacyl-tRNA synthetases and the establishment of the genetic code. *Trends Biochem. Sci.* 16, 1–3.

Schimmel P (2008) Development of tRNA synthetases and connection to genetic code and disease. *Protein Sci.* 17, 1643–1652.

Schimmel P and Ribas De Pouplana L (2000) Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* 25, 207–209.

Schimmel PR and Söll D (1979) Aminoacyl-tRNA synthetases: general features and recognition of transfer RNAs. *Annu. Rev. Biochem.* 48, 601–648.

Schwarz JM, Rodelsperger C, Schuelke M, and Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7, 575–576.

Seburn K, Nangle L, Cox G, Schimmel P, and Burgess R (2006) An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model. *Neuron.* 51, 715–26.

Sgado P, Genovesi S, Kalinovsky A, Zunino G, Macchi F, Allegra M, Murenu E, Provenzano G, Tripathi P, Casarosa S, Joyner A, and Bozzi Y (2013) Loss of GABAergic neurons in the hippocampus and cerebral cortex of Engrailed-2 null mutant mice: Implications for autism spectrum disorders. *Exp. Neurol.* S0014–4886, 34–44.

Shoubridge C, Fullston T, and Gécz J (2010) ARX spectrum disorders: making inroads into the molecular pathology. *Hum. Mutat.* 31, 889–900.

Shoubridge C, Tan MH, Seiboth G, and Gécz J (2012) ARX homeodomain mutations abolish DNA binding and lead to a loss of transcriptional repression. *Hum. Mol. Genet.* 21, 1639–1647.

Siepel A, Pollard K, and Haussler D (2006) Proceedings of the 10th International Conference on Research in Computational Molecular Biology 2006. New methods for detecting lineage-specific selection. *RECOMB 2006: April 2–5, 2006, Venice Lido, Italy* , 190–205.

Snead O (2000) Evidence for a G protein-coupled gamma-hydroxybutyric acid receptor. *J. Neurochem.* 75, 1986–1996.

Snead Or and Gibson K (2005) Gamma-hydroxybutyric acid. *N. Engl. J. Med.* 352, 2721–2732.

Sobreira N, Cirulli E, Avramopoulos D, Wohler E, Oswald G, Stevens E, Ge D, Shianna K, Smith J, Maia J, Gumbs C, Pevsner J, Thomas G, Valle D, Hoover-Fong J *et al.* (2010) Whole-genome sequencing of a single proband together with linkage analysis identifies a Mendelian disease gene. *PLoS Genet.* 6, e1000 991.

Spencer G (2002) Background on Mouse as a Model Organism. *The National Human Genome Research Institute* (NHGRI), http://www.genome.gov/10005 834.

Squire L (2004) Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* 82, 171–177.

Squire L, Stark C, and Clark R (2004) The medial temporal lobe. *Annu Rev Neurosci.* 27, 279–306.

Squire L, Wixted J, and Robert E (2007) Recognition memory and the medial temporal lobe: a new perspective. *Nat. Rev. Neurosci.* 8, 872–883.

Strachan T and Read AP (2003) Human Molecular Genetics 3rd edition. Garland Science.

Swanger S and Bassell G (2011) Making and breaking synapses through local mRNA regulation. *Curr. Opin. Genet. Dev.* 21, 414–421.

Tessier-Lavigne M and Goodman C (1996) The molecular biology of axon guidance. *Science* 274, 1123–1133.

Thomas P, Kejariwal A, Guo N, Mi H, Campbell M, Muruganujan A, and Lazareva-Ulitsky B (2006) Applications for protein sequence function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* 34, W645–W650.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, and Narechania A (2003) PANTHER: A Library of Protein Families and Subfamilies Indexed by Function. *Genome Res.* 13, 2129–2141.

Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, Kaname T, Naritomi K, Kawame H, Wakui K, Fukushima Y, Homma T, Kato M, Hiraki Y, Yamagata T *et al.* (2012) Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* 44, 376–378.

UniProt-Consortium (2011) Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.* 39, D214–219.

Valnegri P, Sala C, and Passafaro M (2012) Synaptic dysfunction and intellectual disability. *Adv. Exp. Med. Biol.* 970, 433–449.

Vardya I, Drasbek KR, Gibson KM, and Jensen K (2010) Plasticity of postsynaptic, but not presynaptic, GABAB receptors in SSADH deficient mice. *Exp. Neurol.* 225, 114–122.

Vazquez LE, Chen HJ, Sokolova I, Knuesel I, and Kennedy MB (2004) SynGAP regulates spine formation. *J. Neurosci.* 24, 8862–8872.

Vengoechea J, Parikh AS, Zhang S, and Tassone F (2012) De novo microduplication of the FMR1 gene in a patient with developmental delay, epilepsy and hyperactivity. *Eur. J. Hum. Genet.* .

Vissers L, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BWM, Hoischen A, de Vries BBA, Brunner HG, and Veltman JA (2010) A de novo paradigm for mental retardation. *Nat. Genet.* 42, 1109–1112.

Vogel K, Pearl P, Theodore W, McCarter R, Jakobs C, and Gibson K (2012) Thirty years beyond discovery – Clinical trials in succinic semialdehyde dehydrogenase deficiency, a disorder of GABA metabolism. *J. Inherit. Metab. Dis.* [Epub ahead of print].

Wagenstaller J, Spranger S, Lorenz-Depiereux B, Kazmierczak B, Nathrath M, Wahl D, Heye B, Glaser D, Liebscher V, Meitinger T, and Strom TM (2007) Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation. *Am. J. Hum. Genet.* 81, 768–779.

Wakasugi K, Slike B, Hood J, Ewalt K, Cheresh D, and Schimmel P (2002a) Induction of angiogenesis by a fragment of human tyrosyl-tRNA synthetase. *J. Biol. Chem.* 277, 20 124–20 126.

Wakasugi K, Slike B, Hood J, Otani A, Ewalt K, Friedlander M, Cheresh D, and Schimmel P (2002b) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proc. Natl. Acad. Sci. USA.* 99, 173–177.

Warden M and Miller E (2010) Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30, 15 801–15 810.

Webber C, Hehir-Kwa JY, Nguyen DQ, de Vries BBA, Veltman JA, and Ponting CP (2009) Forging links between human mental retardation-associated CNVs and mouse gene knockout models. *PLoS Genet.* 5, e1000 531.

Wechsler D (2004) The Wechsler intelligence scale for children–fourth edition. Pearson Assessment, London.

Wechsler D (2008) Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV). Pearson Assessment, London.

WHO (2010) Chapter V – Mental and behavioural disorders – (F00-F99). *International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) Version for 2010* Mental retardation, F70–F79.

WHO (2013) Children's right to family life – Definition: intellectual disability. *World Health Organization Regional Office for Europe* `http://www.euro.who.int/en/what-we-do/health-topics/noncommunicable-diseases/mental-health/news/news/2010/15/childrens-right-to-family-life/definition-intellectual-disability`.

Wirth E, Conrad M, Winterer J, Wozny C, Carlson B, Roth S, Schmitz D, Bornkamm G, Coppola V, Tessarollo L, Schomburg L, Köhrle J, Hatfield D, and Schweizer U (2010) Neuronal selenoprotein expression is required for interneuron development and prevents seizures and neurodegeneration. *FASEB J.* 24, 844–852.

Wolf Y, Aravind L, Grishin N, and Koonin E (1999) Evolution of aminoacyl-tRNA synthetases - analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* 9, 689–710.

Wolfson A and Uhlenbeck O (2002) Modulation of $tRNA^{Ala}$ identity by inorganic pyrophosphatase. *Proc. Natl. Acad. Sci. USA.* 99, 5965–5970.

Wong C, Chan K, Gibson K, and Snead O (2004a) Gamma-hydroxybutyric acid: neurobiology and toxicology of a recreational drug. *Toxicol Rev.* 23, 3–20.

Wong C, Gibson K, and Snead O (2004b) From the street to the brain: neurobiology of the recreational drug gamma-hydroxybutyric acid. *Trends Pharmacol Sci.* 25, 29–34.

Wu Y, Ali S, Ahmadian G, Liu C, Wang Y, Gibson K, Calver A, Francis J, Pangalos M, and Carter O Snead (2004) Gamma-hydroxybutyric acid (GHB) and gamma-aminobutyric acidB receptor (GABABR) binding sites are distinctive from one another: molecular evidence. *Neuropharmacology* 47, 1146–1156.

Xu X, Shi Y, Zhang H, Swindell E, Marshall A, Guo M, Kishi S, and Yang X (2012) Unique domain appended to vertebrate tRNA synthetase is essential for vascular development. *Nat. Commun.* 3, 681.

Yao P and Fox P (2013) Aminoacyl-tRNA synthetases in medicine and disease. *EMBO Mol. Med.* 5, 332–343.

Zhou H and Clapham DE (2009) Mammalian MagT1 and TUSC3 are required for cellular magnesium uptake and vertebrate embryonic development. *Proc. Natl. Acad. Sci. USA.* 106, 15 750–15 755.

Publications and presentations

Publications

Püttmann L, Stehr H, Garshasbi M, Hu H, Kahrizi K, Lipkowitz B, Tzschach A, Najmabadi H, Ropers HH, Musante L, Kuss AW. (2013) A novel *ALDH5A1* mutation is associated with Succinic Semialdehyde Dehydrogenase Deficiency and Severe Intellectual Disability in an Iranian family. *Am. J. Med. Genet. A* Review process benevolently completed.

Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, **Püttmann L**, Vahid LN, Jensen C *et al.* (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 478, 57-63.

Rafiq MA, Kuss AW, **Puettmann L**, Noor A, Ramiah A, Ali G, Hu H, Kerio NA, Xiang Y, Garshasbi M, Khan MA, Ishak GE, Weksberg R, Ullmann R, Tzschach A *et al.* (2011) Mutations in the alpha 1,2-mannosidase gene, *MAN1B1*, cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* 89, 176-182

Kuss AW, Garshasbi M, Kahrizi K, Tzschach A, Behjati F, Darvish H, Abbasi-Moheb L, **Puettmann L**, Zecha A, Weissmann R, Hu H, Mohseni M, Abedini SS, Rajab A, Hertzberg C *et al.* (2011) Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum. Genet.* 129, 141-148

Müller A, **Püttmann L**, Barthel R, Schön M, Lackmann JW, Narberhaus F, Masepohl B. (2010) Relevance of individual Mo-box nucleotides to DNA binding by the related molybdenum-responsive regulators MopA and MopB in *Rhodobacter capsulatus*. *FEMS Microbiol. Lett.* 307, 191-200

Oral Presentations and Posters

Püttmann L, Hu H, Kahrizi K, Garshasbi M, Wienker TF, Stehr H, Tzschach A, Najmabadi H, Kuss AW, Ropers HH, Musante L. (2013) Mutations in *WARS2* and *SARS* are associated with intellectual disability. – *poster* First International GENCODYS Conference, Integrative Networks in Intellectual Disabilities, Paphos, Cyprus

Püttmann L, Hu H, Kahrizi K, Garshasbi M, Tzschach A, Najmabadi H, Musante L, Kuss AW, Ropers HH. (2011) Identification of missense mutations in *SARS* and *ZNF697* in an Iranian Family with ARID. – *talk* 15 th International Workshop on Fragile X and Other Early-Onset Cognitive Disorders, Berlin, Germany

Püttmann L, Garshasbi M , Tzschach A, Kahrizi K, Najmabadi H, Ropers HH, Kuss AW. (2010) Identification of genetic defects underlying autosomal recessive mental retardation and related disorders in Iranian families. – *poster* 23 rd Course in Medical Genetics, EuroMediterranean University Center of Ronzano, Bologna, Italy

Curriculum vitae

For reasons of data protection,the curriculum vitae is not included in the online version.

9 Appendix

9.1 Confocal immunofluorescence microscopy: endogenous SARS localizes to the cytoplasm

To investigate whether the detection of subcellular localization of SARS depends on the cell permeabilization technique, cells were permeabilized with methanol (see section refpermeabilization). Concordant with the findingsfor cells permeabilized with Triton® X-100, confocal microscopy revealed cytopsolic localization of the endogenous SARS protein (Figure 9.1).

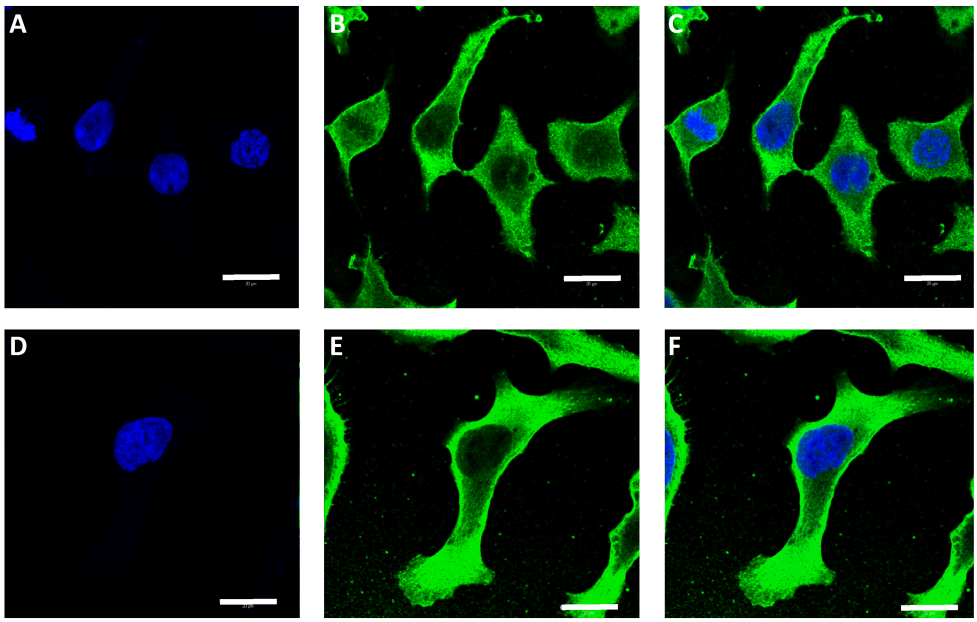

Figure 9.1: Confocal immunofluorescence microscopy showing the cytoplasmic localization of endogenous SARS in HEK293-T **(A–C)** and HeLa **(D–F)** cells. Cells were permeabilized with methanol. The green signal corresponds to SARS staining, whereas the blue corresponds to nuclear DAPI staining. Scale bars = 20 µm.

9.2 Confocal immunofluorescence microscopy: ectopic SARS p.T429A expression in mammalian cell lines

In the study of subcellular localization of ectopic SARS proteins, SARS p.T429A was used as a further control and was found to localize to the cytoplasm in all three cell lines (Figure 9.2).

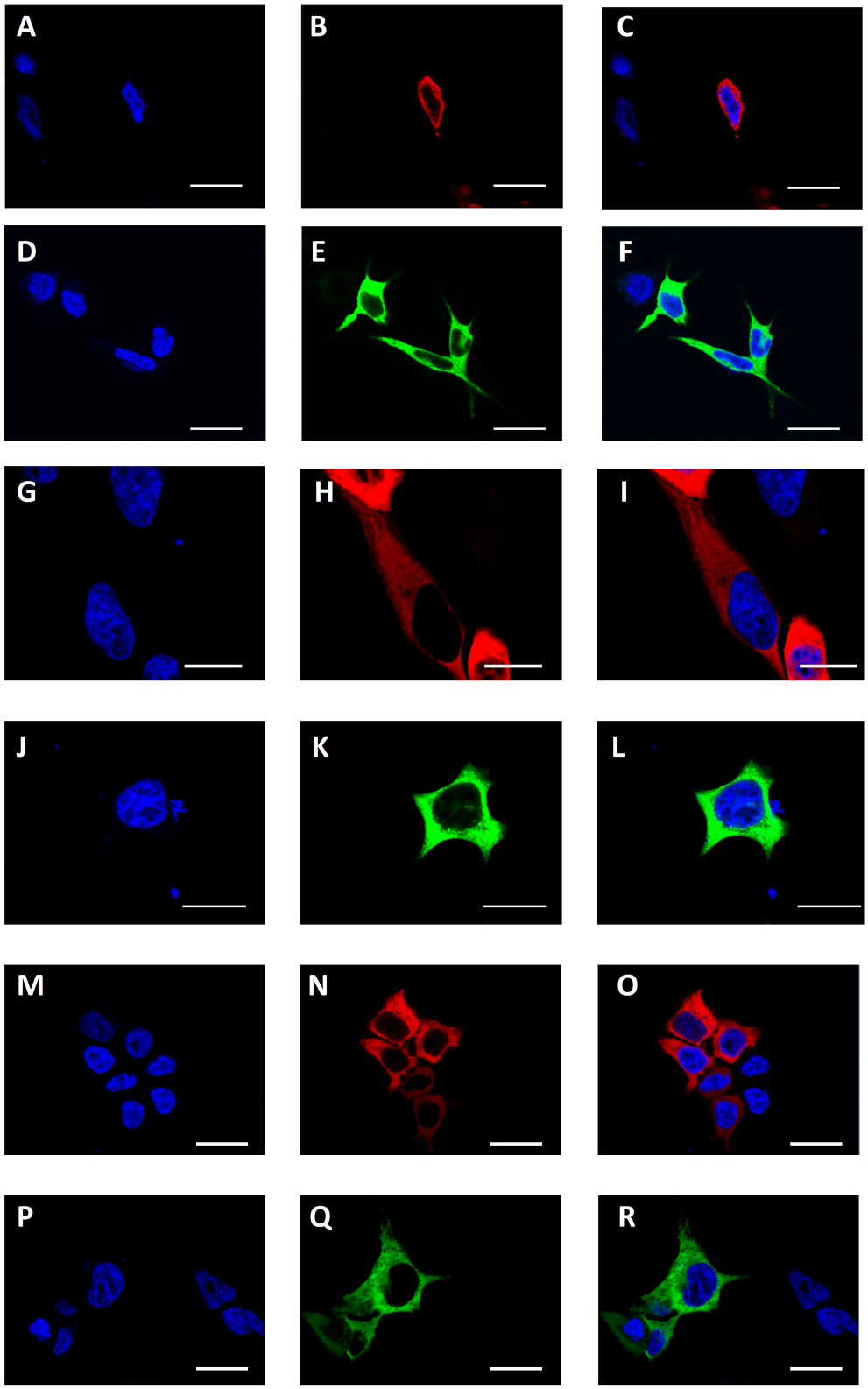

FIGURE 9.2: Confocal immunofluorescence microscopy showing the cytoplasmic localization of p.T429A SARS-mCherry and SARS-EYFP in **(A–F)** SH-SY5Y, **(G–L)** HeLa and **(M–R)** HEK293-T cells. Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

9.3 Confocal immunofluorescence microscopy: ectopic ZNF697 proteins localize to the nucleus

ZNF697 belongs to the C_2H_2 zinc finger superfamily. C_2H_2 zinc fingers induce interactions between DNA and proteins [Miller *et al.*, 1985] and are involved in transcription, translation, metabolism, and signaling [Gamsjaeger *et al.*, 2007; Pabo *et al.*, 2001]. The exact funtion of ZNF697 is currently unknown. To date, predictions about localization of ZNF697 to the nucleus and its involvement in transcription regulation are solely inferred from electronic annotation (`http://www.ebi.ac.uk/QuickGO/GProtein?ac=Q5TEC3`). To gain insight into ZNF697 function, I transfected ZNF697-EYFP and ZNF697-mCherry wild-type and mutant (p.P158T) constructs into three different cell lines. The neuroblastoma cell line SH-SY5Y was chosen to study ZNF697 wild-type and mutant function in a neuronal system (Figure 9.3). As RT-PCR revealed that ZNF697 is well expressed in all cell lines tested, HeLa and HEK293-T cells were chosen as basic model system (Figures 9.4 and 9.5).

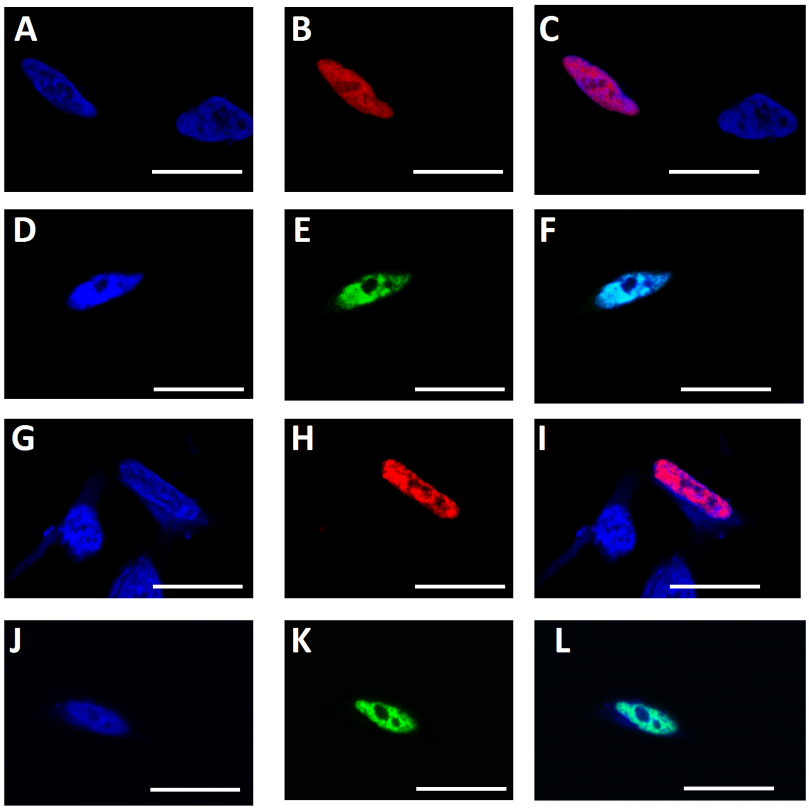

Figure 9.3: SH-SY5Y cell line: nuclear localization of EYFP-ZNF697 and ZNF697-mCherry **(A–F)** wild-type and **(G–L)** p.P158T. Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

The prediction about nuclear localization of ZNF697 could be confirmed in all cell types studied and neither tag interfered with the transport of ectopic ZNF697 to the nucleus.

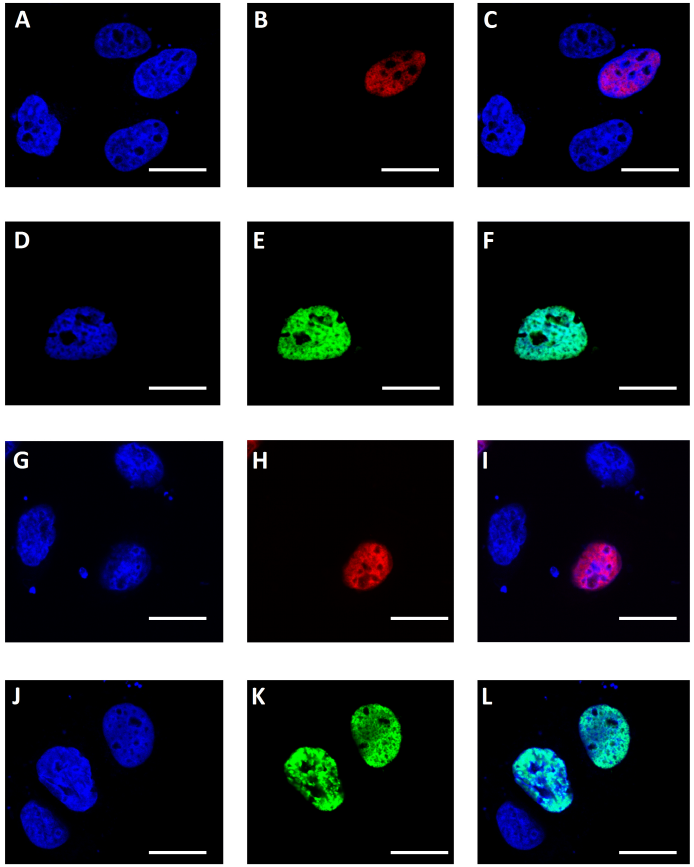

Figure 9.4: HeLa cell line: nuclear localization of EYFP-ZNF697 and ZNF697-mCherry **(A–F)** wild-type and **(G–L)** p.P158T. Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

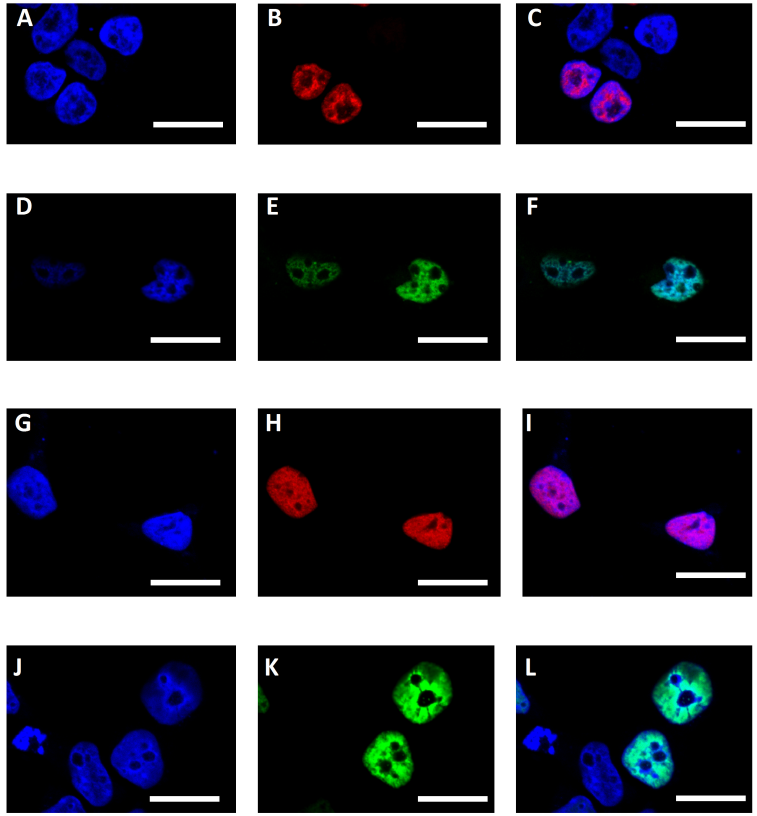

Figure 9.5: HEK293-T cell line: overexpression of EYFP-ZNF697 and ZNF697-mCherry **(A–F)** wild-type and **(G–L)** p.P158T . Green: EYFP fluorescence; red: mCherry fluorescence; blue: DAPI. Scale bars = 20 µm.

9.4 Cell fractionation: endogenous SARS localizes to the cytoplasm and the nucleus

To investigate whether endogenous and ectopic SARS localize to the cytoplasm, cellular fractionation analysis was perfomed with HEK293-T cells transfected with *EYFP-SARS* as well as with untransfected cells. Endogenous and ectopic SARS were present in the nuclear and cytosolic fraction (Figure 9.6) (see section 4.11).

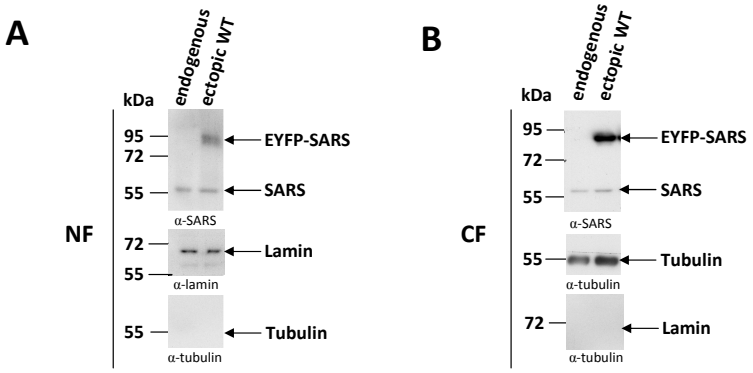

FIGURE 9.6: Expression of endogenous and ectopic SARS proteins. **A)** Nuclear (NF) and **B)** cytoplasmic (CF) fractions from HEK293-T cells either untransfected or transfected with wild-type EYFP-SARS were run on SDS-PAGE. The gel was blotted and probed with α-SARS antibody. The blot was subsequently probed with α-tubulin and α-lamin as loading controls for cytosolic and nuclear fractions, respectively.

9.5 Primers used to sequence the protein coding genes located in the linkage interval identified for family M289

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
ADAM30_ex1.1	CCCCAGAAATCTGACTCGC	AGCCCATGTAGTTGCAGTCC	MAGI3_ex1.1	GTCCAAGGGGCGTCTCG	AACACCCCGAGACAGATACG
ADAM30_ex1.2	TGTTGCCCCGACATCTG	TTGTTCACAAACCTATACCTACTTTG	MAGI3_ex1.2	CAGGAGTGCGCCGTGTC	AAGATTACCCGGGACAATACG
ADAM30_ex1.3	ATAGAATGGCAGATGGCCC	GAGAACACACTTTTCCAAACGAC	MAGI3_ex2	CCTGTTTATAGCTAACTGGTTTAAGAG	TCTTGTTCCCAACTTAGTCTTCAC
ADAM30_ex1.4	CCAGAGTTAGCTGAAGTTTTAGGC	AACACCGATCTTTCTGACACTC	MAGI3_ex3	GCAGGGGAATGAAACATCTC	TTTACTGGTCAAATTCCCTCTG
ADAM30_ex1.5	TCAGGACGCACTGGGTTTAG	CCTCCATGGCATCAGGTC	MAGI3_ex4	TGCTCTGACTCTCACTGCAAC	AAAATTAAAATCCCATCACTTGC
ADAM30_ex1.6	TCAAGTTCCTGCCCAAATG	CAGGAGGTGCCATCATTTATC	MAGI3_ex5	GGAAAGGTGCTGTTTGAAGAAG	TCAAATACACTATTCCTTTATCCAAG
ADAM30_ex1.7	TGAAACCATCCCTGATTTGC	CCTGTTTGGGTTTTAAGTGG	MAGI3_ex6	CAAATGGAGCCTGTTGGAG	AAACCCATTATCTGAGCAAGAAG
ADAM30_ex1.8	GCTCAGAGGGGCGATTC	GAAGAATGGAAAGCCTTTGG	MAGI3_ex7	CACTCACTGGTGAATTCTTTTCTG	AACAGCTGGATCACAGAAGC
			MAGI3_ex8	GCAAAGTAAATGTTCAACTTGTGTG	AGGTCTAATTCAGAAGCAATGG
ADORA3-iso1_ex1	AGACGGATCTTGCTGGCTC	TTGTAGCCTCATTTCCCAGC	MAGI3_ex9	GCCTAAAGATACATATTTGTCCCC	GACACTCTTTAGTGTCAAAGCCATAG
ADORA3-iso1_ex2	TCTTCAAACCAGGATAAAGCC	TGCAAAGTCTCTGGAAAGGC	MAGI3_ex10	GCACAGACAAGGCAATGAGA	ACCTACTGTATACCCATACAACTTTTT
ADORA3-iso1_ex3	GGAGCTGCAATGTTTTGAGAG	CTCTCTTAGCAGCAAACACCC	MAGI3_ex11.1	TTTCCAAAATTCAAATTTAGTTGC	TGTGTCCCGATTTTCCATTC
ADORA3-iso1_ex4	GCGTGGTACCTCTGCTTTAC	TTTTTGAGGAAACCAAATACTCAG	MAGI3_ex11.2	CATTGTTGCTGCTACCCCTG	AAGGCCTTGACACCACTGAC
ADORA3-iso1_ex5	TCCAGAGCTGAGAATCAATAGG	TTGAAATCTGCTCCTTGACC	MAGI3_ex11.3	GTATCCATGGCATCGTCAGG	CATTCTAAACTGTAAGGAAAGACAATC
ADORA3-iso1_ex6	TTCCTACTGATGGGCTTTGC	CTGGCTGCAAATGATTGTTG	MAGI3_ex12	GCAGTATTTGTGATCTGGGAATG	GGGAATGTGGCATCTCTCTC
ADORA3-iso2_ex2.1	CTTTACCAGCTCCATCTCCC	CAGAAACAAGGACTTAGCCG	MAGI3_ex13	TGTCTTTGATCCTGCCACC	TCTGAAAGATGGGAAGGTCC
ADORA3-iso2_ex2.2	TGTGCGCCATCTATCTTGAC	CCCACCTCAGTGGAAGTAATC	MAGI3_ex14	TGTGTGTGTTACACTCTTGTGC	CACCACAATAACTGAGGGACTATG
ADORA3-iso3_ex2	CAGATTCAGAATTCCTGGGC	TGCAAAGTCTCTGGAAAGGC	MAGI3_ex15	TCTGTTGGTTTGCACAATGC	AACCTGCAAAACTACACAGGAAG
ADORA3-iso3_ex5	CCAGAGCTGAGAATCAATAGGC	TTGAAATCTGCTCCTTGACC	MAGI3_ex16	CTCACATGTGGGGCTGC	GGGAGCCTCAGAAAACTGG
			MAGI3_ex17	TGAAAATAGTTCCGTTTGGATG	ATATCCCCGCAGTGATTCTG
AHCYL1_ex1	AGAAGCCGACGCAGCTC	GTCCCAAACTAGCCAGAGCC	MAGI3_ex18	ATCCGCTTGGCCTTCAC	TGGGGTTATGAACAGCAGAG
AHCYL1_ex2	TGGCTTTTAGGGAAAGACAATG	CACTGGTAGCAAACAGGGAAG	MAGI3_ex19	GGAAAATGGCAAGGTCGAG	AATGCCAGTGGTATGGCAAC
AHCYL1_ex3	TTGGGGACTGCACTCTGTAG	CCCAGTACACTTTCAAGTCAGC	MAGI3_ex20	TTTCTTCAAAATGTAAGTAAACAGTCC	TAGGAGGTTCCTGCATTCCC

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
AHCYL1_ex4	GGGGCAGGAACTGATGATAC	GGCGAGTTAGGAAAGGAAAAG	MAGI3_ex21	ATGGCAGTACTGGGTACACG	TTTCATGGTTCTAGCATTCACC
AHCYL1_ex5	TTCCTGGGTGGCCCTTC	TCCACTTTCTTCCTGTTGAGG	MAGI3_ex22	CCGCTTTTGAGGGATAAGTG	TGTCAAATTTAATGTCAAAAGAAGAC
AHCYL1_ex6	AGTGATTTGGTACAGGACACTCC	TAGGGAACCATGCTTTGAGG			
AHCYL1_ex7	AAGTGGGTTCAAAGTTCCCC	GAGACCAGGAGAGTTTGCTGAG	MAN1A2_ex1	CGTGAATGACGTGCCCTC	ATCCAATCACCGTACCTTGC
AHCYL1_ex8	GAATTAGGCAGATTTTGCCAC	AAGCCAAAGGAAGACCAGAG	MAN1A2_ex2	GGCAAAAGTAGAAAACAAGTTCC	CAACTATGACCCAAGAATAAACAG
AHCYL1_ex9	TTGTGAAGAGGGACTGAGTTTG	CAGCACAAGCCTCCTTTACC	MAN1A2_ex3	GGAAGAATGGGCAAGGATATAG	AACTTAATCGTGACTGACACTTGC
AHCYL1_ex10	CCTACCTACCCTAAATATCCATGAC	TGAACTCTCAATAACATAAAGCACC	MAN1A2_ex4	AAATACACTGAATGTGAAGGGAATC	AGGAAGGCACAAATGTTCAG
AHCYL1_ex11	ACCATCTTCCCTCCCACTG	CCTCCTTTGTATGAAAGTCACTC	MAN1A2_ex5	TTTTCCACAGGGTGTTTGTG	ATTAAATTCCAACCCTCCCC
AHCYL1_ex12-13	CAGCCAGTACCTTTGTTGCC	TTTGTTTCCCTCCCCTCC	MAN1A2_ex6	TCTCTGATTGCCAAATGCTG	GACAATTAACCTGTTTCAAAGGAC
AHCYL1_ex14	GCTTTGGCTTGACTTGTATGC	ACTGGAACTGATCTACGCCC	MAN1A2_ex7	CGAATGTTTTGGTCTTTTCCC	TTGGGCCAATCTTAAAGGAC
AHCYL1_ex15	CTGTGACTGTGGTAGATCACTGTAG	AGCACAGGAACAGGAATTGG	MAN1A2_ex8	AGCCTACATTAAAACCAAGCC	CAACATCAAGCAAAGAGTTTTCAG
AHCYL1_ex16	AAAAGCTTCAAATGAAAATGGG	GCCTGCTTGGTTTAATACTGG	MAN1A2_ex9	TTGTCTTGTCTTCCCTTGCC	CGTGAGCCACTGCACCTG
AHCYL1_ex17	AATAAGGGAGACTGGTGCCC	GTGTGGTCTGCAGAACTTGG	MAN1A2_ex10	GGTTTTATAATGGCCAGAAGG	TGCCTATATATTCTCCTTCTTATTGC
			MAN1A2_ex11	CCTTGCAAACAAACATAAATGG	ACATGGGTGAGGGTATGTGG
ALX3-1-FW	GGATGGTTCCAGCATTAAGTCAGAG	AGGGGCAAAAAGTTGAGAAATAGGC	MAN1A2_ex12	TTGTCCTTCAAAGTTTTGTTGTC	GATCACACAGCAAAACCAGTC
ALX3_ex2	CCATTCCAACCCACAAAAG	AGCCCTTCCAGATCACTTTC	MAN1A2_ex13	TGGTTTTCTGTGCCATTAAAAG	CATGTTGACATAGAGGTCCAGG
ALX3_ex3	TAGCAGGCTCCTTTCCTCAG	TTCAGGCCAGACCTCATATTC			
ALX3_ex4	ACAAAGAGAAAGCCCCTGG	GGAGCGACTGGGAATGG	Mov10_ex2	GGATCAGGGGTCAGAGTTAGG	CAGAACTGGCGGGCTCC
			Mov10_ex3	GATGCCCTTTCCCACCC	CCCTTCCCTCACACCCTG
AMIGO1_ex2.1	CCCACCAGAACATGCCC	TGTAGAGCAGCAGCACCTCC	Mov10_ex4	CAGGGTAACTCCCCAGTCAG	TGTTTCCCAAAGCATAAGCAC
AMIGO1_ex2.2	TCTCCTCTGAGGCCTTTTCC	GACATTGTGCAGCTTCTTGG	Mov10_ex5	GGGATAAGGATATGGGTGGG	GCTCTGTGCTCTTCCCAAAG
AMIGO1_ex2.3	CCCTGAACTGCGACTGTGAG	GGACCAGGACCACACTAAGG	Mov10_ex6	CCAGAGGGGCGTTCTCAG	GCAATGCAAAGAAGCCAGTAG
AMIGO1_ex2.4	GGACGGTGGTGTGTATACCTG	CTCATATCCTTCAGGGGTGC	Mov10_ex7	ATAGGGGTTGTCTGGCCC	GGAATAGTTCTGCCCACAGC
			Mov10_ex8	AAGGTCTGGGGAACCTGG	CTCCCCTGGGGTCAGAAG
AMPD1_ex1	TCACCCCACAGTCTCCTCTC	TGAATGATCAAAGCTGATATGGTAG	Mov10_ex9	AGTTCTGGCCCTTTATTGCC	AGGCCTCCCTGACTCCAC
AMPD1_ex2	AGACAAGCTGGAAGGCTGAG	TGAATCTTTGTTGACTTTCAAGG	Mov10_ex10	AGTCAGGGAGGCCTCTGG	AATCTGGGGACAACTGCTTG
AMPD1_ex3	TCCCATGCCTTACACATAGC	TGGCAGATACCCCTCCTTAG	Mov10_ex11	CCTGGGTACGGCTGAGTC	CTAGCTTCCTCGCCTCAGTG
AMPD1_ex4	GGATCCTGTTTCTGATACCAAC	TGAAGTGTAACAAAGGACAGGTG	Mov10_ex12-13	CAGAATCACGGTGGGAATG	TGCAACCTCTCTGACCCAC
AMPD1_ex5	TACACAGGAGAGCTGGGAGG	TGGGGCCAAAGATGATTATG	Mov10_ex14	TGTGTGGGTCAGAGAGGTTG	GCATGCAGGATGTGAAAGC
AMPD1_ex6	TACCGGATTCCAAGCTAATG	TGAAGGGCTTAATACAGTGTTTG	Mov10_ex15	TGTAGGGCAGAGGGAATCTG	TTCAGTGGGAGGGAAGTCAG
AMPD1_ex7	AATGATAAAGGTCCTGATAGTTTTG	AAATCCACTGAACCTGTAGAAATC	Mov10_ex16	TTTGGGAAGGGAGGGAATAG	TAAGGAGGTACCATCTGGGG
AMPD1_ex8	GATTTGCAGCACCATATCCTC	AAAGACATGTGGGGTTCTGAG	Mov10_ex17	AGCTCCCTGGCCTCCTG	CTGGAAGTGGGCAAGTGG
AMPD1_ex9	AAATTCCATTCTTGGAGCCC	GGATCGCTCCCTCCCTC	Mov10_ex18-19	GTAGGGCACCCCTTTGACAC	CCTAGGCAGGGGCTGTG
AMPD1_ex10	GAACCCCAGAGGAACTTGTG	CAGGGGACTTGAAGCAGATG	Mov10_ex20-21	GGACAGGACCGTGGCTTAG	GTTCAGCTGGGTTCTGGTTC
AMPD1_ex11	TCTGGTTCCCTCATACTGGG	ATGACCAGGTAGGAAGGCTG			
AMPD1_ex12	ACTGGGACCCTGAAGTTGG	CCCCTGACCACCTTAATTGC	MYBPHL_ex1	TCCCTCATCACCAGCAGTG	ATGTCTTCCCCAACTCCCAC
AMPD1_ex13	TTGTGAATCACTCCATACCCAG	TGACAGTATGGTTGGTTAAGGG	MYBPHL_ex2	CAGCTGGATGCAGGATTTAC	TCTGTTTCCAAATCCAAAACC
AMPD1_ex14	GGCAGACCATGAGATTGTAGTG	GACGGGAAACAACAGTCCAG	MYBPHL_ex3	CAGCAGGGAGAGTCAGGAAG	CAGCCTGGAGAGGGAGTTC
AMPD1_ex15	GACTGTTGTTTCCCGTCCAG	AACTGCATATGATTCGTGGAAC	MYBPHL_ex4	GTGACTTCAATGCATGTGGG	GAGCTACCAGAGGCCTGTCC
AMPD1_ex16	TGCAATTATTGTCTATACCAGGG	GGTTTATAGTAGGGACTCAATAACAGG	MYBPHL_ex5	GAAGAGGGCATGTTGGGAC	ACTGTGCCCGGCCTATC
			MYBPHL_ex6	GATAAAGAGTCGACGAGGGC	GGACAGGCAGGAACAAGAAG
AMPD2_ex1	GAGGCAGGGGAGGGATAAG	AGCACCTTGGGCGAGAG	MYBPHL_ex7	GACTATTGGTTCTTACCCTCGC	TGTTCCCTGACTCTTCCACC
AMPD2_ex2	GACCCTGGGCTCTCTCG	AGGGGACAGCTTGAGGTGAG	MYBPHL_ex8	CTCCCCAAGAACCTGAAATAG	TCCCTGACTCAACAACACTAGC
AMPD2_ex3	AGGTACCCCTGGTCCTGC	AGAAAAGGGGCTCTGAAAGG			
AMPD2_ex4-5	CTAGTGCGGTGCTGGGAC	GGGGAACCCACTAAGACATAGG	NBPF7_ex1	TGATCACGTTTTTCTCAACAGT	TTGCTACCTCTGTCTTCCAAA
AMPD2_ex6	GGCCTTAGGGAGGAGGTC	AGGCACCAGGCTGAGAAG	NBPF7_ex2	TGGAAGTCCCCATCTAGACC	GGGTAAGTGGGGTGGTTATG
AMPD2_ex7	CCTCTGTGGGGCCTTTG	AAGCCAACGCAACCTCC	NBPF7_ex3	CTTTCCCTTGGCCACAGAC	TGGCCACATATGTTTAGTGGA
AMPD2_ex8	GCTTGGGAGAGGCCACAG	GAGTCTGGGGCAGAACTGG	NBPF7_ex4	GAAACCATGCCAGGTCATTT	TCAGTGCACTCTGCAGCTTT
AMPD2_ex9	GATCTGCTACCCAAGCCCTC	GCTGGGCAGAGGTGTGG	NBPF7_ex5	CCCTCTTAAAGGGAACCTCCT	CCTCCCAAAGTGCTGTGATT
AMPD2_ex10	AGGCCCCAGACCTTCCTG	ATGGGAAGGCTCAGGTCAC	NBPF7_ex6	GGATCTGAAAAGCAGGCTCA	CCACAATCAGGAAGACAGCA
AMPD2_ex11	TGAGTCAGTCAGGGGACCAG	ACCCTCCTTGAGTCTCACCC	NBPF7_ex7	AGGGCAGTCACACTCCACTC	TCACATTACCCAGCAATGACA
AMPD2_ex12	TCAAGGAGGGTAGGCAGATG	CCCTGACCCCTTGAATATGAC	NBPF7_ex8	AGCAATCCCTCCAAACTCAG	TGGAAATAGGATGAATTGTGGTT
AMPD2_ex13	AAGCTCTGACCCTTGCTGG	TGCAGCTGGATGCGAAG			
AMPD2_ex14	CAGAGCTGGGTATGGGGAG	AAGAGTGAGGTTTGTTGGGC	NGF_ex3.1	CCCAGAAACTGCCTTTTGAC	ACACCGAGAATTCGCCC
AMPD2_ex15	GGTGAGCCCAGGTGATCC	AGGAGGATCGGTTGCTGAAG	NGF_ex3.2	ACTTCGAGGTCGGTGGTG	ACAGGTTGAGGTAGGGAGGG
AMPD2_ex16	AGAACCACTGTGCCCTGC	CACCCAACCCGGGAGAC			
AMPD2_ex17	AGAGTGAGTGAGGAGGCTGG	CAAATGCAGAGGAGGGGAG	NHLH2_ex3.1	CCTTCCAGCAAAATACCCTG	GCGAAGGCCAAGTTGAAG
AMPD2_ex18	GTCCCTGGGATGGCTTG	GGTCTGAGGACAAAACGTGG	NHLH2_ex3.2	GTGTCGGACCTGGAGCC	CCGGATCCGTAGCGTTTC
AP4B1_ex2	GAAGGAGCCCTCGAGTTACC	CCTTTCAGAGAGACTGGGGC	NOTCH2_ex1	GGGGAGTCGAGGCATTTG	TGTCCAAACTCTTGGGAACC
AP4B1_ex3	TCCCTCCATTGTAAGTTGGG	TTTTCCCCAGTATTAAGGAGC	NOTCH2_ex2	GCTGGAGATACTAAAACACAGAGAAA	AGCCAATTATCTGGCATCC
AP4B1_ex4	TCTGTTAGGTGAACAGGGGC	TTTCTACCAAAGCCACACAATC	NOTCH2_ex3	GCTGTTACACCATCCCAACC	ATGGGGTCTTGCTGTGTTTC
AP4B1_ex5	GGAGCTATTACTCTGGCTACATCC	CAAGATCCCACAAGGTAGCAG	NOTCH2_ex4	TGCTGAGGTCCTTGGAGAGT	TCTTTTCCATTCCAGACTTGC
AP4B1_ex6.1	CTGGTCATATGGCAGTGGC	TGGCTGCTAAAGTGACCTGG	NOTCH2_ex5	AATAACCATGTTCTGTGGGAAC	GCAGGCCTAAGATATTTGTTACTG
AP4B1_ex6.2	CAAACTGATGTCCTTGTGCG	ACCACAGTTTTCTTGCTGGC	NOTCH2_ex6	GGCTGGTATGGTACTAGTCCTTTG	GAAGGCAGAGTCCTGAATGC
AP4B1_ex7	CTTTTGTCTTTCACTGGTGGG	TGATGCAAGGAAAAGTTCAAATC	NOTCH2_ex7	GGATTCCGGATAAAATATAGG	CAAAGCAGTTGAGTTTCCAC
AP4B1_ex8	AAGATTCCTGCCTACCTGGAG	GCAGGGTTTCAAGCATTCTC	NOTCH2_ex8	CTGGCCCTGATGCTATAACC	GAATGCTCTACCAAGAGAAGTTCAG
AP4B1_ex9	AAAAGCAAAACCAGTGGCTG	GACGTGACTCACCAAACAATG	NOTCH2_ex9	TGCAAATTTGTGCTTCCTTG	GTGGTCAGTTGGCTTTGTGC
AP4B1_ex10	TGGCACAAGAGTAGGAGGAAAG	TTCCACCAATCATTGAAAAGG	NOTCH2_ex10	TGACCCATTGGTATAAGCCC	GGAGTGGACTGGCCTGG
AP4B1_ex11.1	TTGTGTACATTTGGGAATTTGG	ATCCTGAGCACTGAGGTATGC	NOTCH2_ex11	GCTACACTTGTGTTCTGGTG	CAGAGCTATTGTCTCTGACTTC
AP4B1_ex11.2	ACACCCTCCAGATGGCTC	ATCTGGACTTACTGGCAGCTC	NOTCH2_ex12	CATCACAAGCAGAAAGCCTTAG	TTCCTTTGGATGCTATATTCCC
			NOTCH2_ex13	GCCTGAGCAGATGAAGGATG	CCCCATTTGACAACTTCAGG
ATP1P1_ex1	CACTTGGCTCCCCTGGTC	CTCCCGACTTCCTCCTCTTC	NOTCH2_ex14	CCCATTTCTTCACCAAGTCC	GGAAACTAAGAAGTGAGTCAAGCAG
ATP1P1_ex2	GCCATTCTGAAGCCTACCG	AGAATACAGTGAGGCCTGGG	NOTCH2_ex15	GGTTAGAAATTGATGAGGCTGTC	TCAGGCACTGACAAAGCAAG
ATP1P1_ex3	GCACTGAAGAAAACATCTGCAC	CCCCTGAGGTTAGACCTGTATG	NOTCH2_ex16	GAATTCAGAGAACATCCTTGGG	GCCTCATAAGACCAGCAGGG
ATP1P1_ex4	TATATTGCCTTGTAAGTGCTGG	GAAGTGGGAGACAAAGACGG	NOTCH2_ex17	AGCAACTTCCTAGGCCACAG	GCTCCCTGTGTGTTTTCCTC

Name	Forward primer	Reverse primer
ATP1P1_ex5	TTCCTGTGTGGTCTTTAGAAGG	TGGGTGGGGTAGAAGACATC
ATP1P1_ex6	GGCGAGCAAGCTTTTGTAAC	ATTCACCAGTGAGCGAGGAG
ATP1P1_ex7	AAATGGCTGCAAGGTAGCTC	ACACCTCTGCTCTGTGCTTG
ATP1P1_ex8	CGTGGCTTCCTTCAGGTTAG	AGAGTGTAACATTCGTGCAAGC
ATP1P1_ex9	AATTCCAGCCTGGTTGAGTG	AAACCAAGCGAAGTCAGCAG
ATP1P1_ex10	TCACTGCAGATTTCTATGGGAC	CAGTCACACAGACCACAAAGG
ATP1P1_ex11	CCTCTGACAAGATTGGAATGTG	GTACCTCCCCTCTTTGTCCC
ATP1P1_ex12	TTGTTTTCCACCATGGACTG	AACTGTCTACCAGACTAGGTTACTGG
ATP1P1_ex13	GGCTGAATAGGCTGCTGTTG	AGTAGGGTGTCCCAAGGC
ATP1P1_ex14	TCCTTAGTATTTACTCTTGCCTACTGG	CCATTATAGAACCACGCTGAAG
ATP1P1_ex15	ATGGACACATGCTGGTGATG	CCACTGGAACTCCAATCCC
ATP1P1_ex16	TCGCTAGGAAAAGTGATTGG	CCAAGAAGTATACAGAAATGGATGC
ATP1P1_ex17	TCAGTTTCCAGTGTGCTTGTC	CATGGATGCTGGGAGCATAG
ATP1P1_ex18	CAAAAGGTTCACAATATTAGCTTCC	GACACCATCACCTCAGATGC
ATP1P1_ex19	CAGGTACCCTTGTGTGTGAGG	TTGAGGAGCTCTGACACTGG
ATP1P1_ex20	TCAAGCCCCAGAAGACTGAG	GCAAAGGGTTTACAAATCACC
ATP1P1_ex21	CATCTGACCTCCAAGTTTTCAAG	TTCCAACCCATAGGAATTTTTG
ATP1P1_ex22	GGTGATGTGCAATTTCTCTTTC	CCAATAGAACTATAGCCCCTCC
ATP1P1_ex23	TTCATCCCTGTTTCCCTCTG	GGGTAGAGTTCCAAACTCCAAG
ATP5F1_ex1-2	CTGACAGATTCTCCTATCGGG	GCAGGAATCGGGAAATAAAAG
ATP5F1_ex3	TGTTTCTGCTCATTATTCATACACTG	AACTTGATTTTAGTGGTGCCATC
ATP5F1_ex4	AAGTCATTACTTTGCCTTCAGC	TTGCTCCCCTAACACTCATATC
ATP5F1_ex5	AAGAAGCATGAAATGAATCTCC	CGTGAATGCCAACTCCTACAC
ATP5F1_ex6	CGTAAAATACAAAAGAAGTGGGC	GCTAAAGACTCCTTGCCAGC
ATP5F1_ex7	TGAGGTTCTGGTCTGGTAGGA	GGCAGAGCAAGAAGGATGAG
ATXN7L2_ex1	CAAAGACTGCGCTGGAAAG	GCGGGTCCTTTGTTTGG
ATXN7L2_ex2	CTCTGGTAAGGAGAGGCTGG	GGAAAAGTGGAGCGTGGTC
ATXN7L2_ex3	AGTGCTCTCCGATCTTCCAG	GAGTTCCATGCTCCACCTTC
ATXN7L2_ex4	GGTTAGGTTGGGGAAGGAAG	TGTATTGATGTCAGCCCATCC
ATXN7L2_ex5	AGAAGATGTCCAGTCTCCCG	AGGACAAAGCAGCCCAAAG
ATXN7L2_ex6	TCTCAGGCCCTTGGCAG	AGTGGGCAAGCCCTGTG
ATXN7L2_ex7	GTCCCTCACCCCTTCTCAG	CTGAGCCTGTTTCCAGGG
ATXN7L2_ex8	CCCTACTCTGACCCCACAAC	ACACAGGATCTGGCCTCCTC
ATXN7L2_ex9	CCAGATCCTGTGTGCAGATG	GCCCAAGTATGTCTGGAACC
ATXN7L2_ex10.1	TGGAGTGTGTCTTGTGTTGG	CTTGCCTTTGGATGACTTGC
ATXN7L2_ex10.2	ACAACCTTGTCCCCAGCTAC	AAGGTTCACCAGGCAAAAAG
ATXN7L2_ex11	CACAGTTCACTGGGTAGGACAG	ATCTGGAGGGGAGGAGGTG
BCAS2_ex1-2	TGATTACATCAGTGTCAGGCG	GCAGGCAATGGTTAAAACTG
BCAS2_ex3	TGGGTAGTGTAGCTTAAATCTGTTC	AGCCAGCAAATACATTCTTCAC
BCAS2_ex4	AGAGATGTAAAGGCTTGGCAG	AAACTTTGGGTGAGATGGAAAG
BCAS2_ex5	TCCAACATAGTTCTTGGCACAG	ATCTGGCTGCAGTCTCTTTG
BCAS2_ex6	TGTTCTCATGGGATTAAGATATTCTC	TACCTGTGGCATCCCATTTG
BCAS2_ex7	GCAACAAGAGCGAGACTCC	GATTTCTAAACACTTAAACATCAATGG
BCL2L15_ex1	TTTAACTTCCTGCTGGGCTG	TGTTACTGAAAGCTTAGAGTTTGCC
BCL2L15_ex2	ATTCCTGGGACTTGCCATTG	TTTTCATAGAACATTCCTAGTCACTG
BCL2L15_ex3	TGCTTTACTGGATTCCTGCAC	TCATGAATTGGAATCCCAGG
BCL2L15_ex4	TCTGCAGATCATTTCCTGG	CACCCAATCAGTCAACAAGG
C1orf59_ex2	GCTGATTCTCTCAAATAGAGCTTG	GCAATTATCTCACAATACATCCAGG
C1orf59_ex3	GCCCCTTATTTAATTTCTGAAGTC	CCCCTGTGAAAAGTCTCTCC
C1orf59_ex4	TGTGTGTAGCATCTTCTGCC	TGAGTTGAAGATCCCTAATGGC
C1orf59_ex5	CAGAGGATCTGAGTTCTGTAGCTG	CAACCAATGAGCAGTTTTCC
C1orf59_ex6	GGAGAAATCGATACACACAGATGG	TTACATACTAAGTCACACATTGGAATC
C1orf59_ex7	CCAGAAATGGTTCCTTATTTAACAG	TCATTTACACTCATCTCTACTTTGGG
C1orf59_ex8.1	CCCATTATAACTGAGGCATTTACG	CTTGGGATACGCAAGGAGTC
C1orf59_ex8.2	GGGGTGATAAGCCCAAAGAC	GGAACATAGACTTTTGCAGTGAAAC
C1orf62 is *AKNAD1* in hg19		
C1orf62_ex2.1	ACTGACAGCAGCCAGCTCAG	TGCTGGAATATGAAGAGCTGC
C1orf62_ex2.2	CAGCTGTGACCATACCCCTG	ATTGCTGTTTTCACCATCCC
C1orf62_ex2.3	TGCCGACAGTTTTGAAGAGG	TTTTCACTTTGGGAGCAATC
C1orf62_ex2.4	AGTCACCCCAGAAACAGCAG	CCCACACTGAAAGAGATAAGCAG
C1orf62_ex3	TGGGCTCCCTCTATGATGAC	TTGTGCTGCTGCACTCC
C1orf62_ex4	TGGTGACTGCAACAGCTTTC	TCTCTCATTAATGGCATGGTTTC
C1orf62_ex5	GAAGTTGAAAGAACAGACTGATCAAC	AGCAAAACATTTCCACTCCC
C1orf62_ex6	TGGGAAAACAATTACACACACAG	CCACATCAGAATCCCACCTG
C1orf62_ex7	CAGCAGGTGAAGATCAGAGC	AAAATAATATAATTCATGGGACTTTGG
C1orf62_ex8	TTTGAGTCACCTCCAGTGGG	GAAGGGGCCTCTTTCTGTG
C1orf62_ex9	CAAATAGTGGAGCTCCCAGG	CACTGGTTTTAGTTAAAGAGCCAAC
C1orf62_ex10	AAGTTCAGCCGGACCTCAG	CAGCCAGTGGCCTCAGTTAG

Name	Forward primer	Reverse primer
NOTCH2_ex18	CCCCAGCATGACTTAGTTCAG	CCCTGCTCCACAATTCTAGC
NOTCH2_ex19	GCCTCTGTATGATCTTAGATTCCAG	TCTCTTTTAGAAGGAACTTATTCAGAC
NOTCH2_ex20	AAAATGGTTAGGTGTACCTCGG	GCCCACCCACTACTATCTGC
NOTCH2_ex21	TGATTTGAACCCCATGTGTC	AAATTCAATATATCAGTGCTCAAACAG
NOTCH2_ex22	GAGGTGGGAGACAGTCACAG	AAATGCTTGGAATGAAGCTAGTC
NOTCH2_ex23	TATCCCTTGCTGGAAAGCAG	ACAGGGCCCAATTTGTTCAC
NOTCH2_ex24	TGAGTAACTTGCAGTGACTACTTCTG	CCAGGTTTAATCTCAGTTCTGTTC
NOTCH2_ex25.1	GAACGGGAAGTACAGAAGGG	GTTGCAGGCCTCATCACAG
NOTCH2_ex25.2	AGCGCCAGCCTCCTTATTAC	GCACCATCTGAAAAGCAGAG
NOTCH2_ex26	AAATGGGTAGGAAAATCCACAG	GCTATATGTCAAAGTGCTAGGCTTC
NOTCH2_ex27	GACATGTTCTGCCTGACCTG	CCCCAATTGACACTCTTCCC
NOTCH2_ex28	GAATCTAATGCTGACATTGAGAGG	CAAGATATGCTTTTCTAGTCATCCC
NOTCH2_ex29	GGTGGAAAGTGTTGAAAACCAC	AGGACAGAAATTTGAAATGTAACC
NOTCH2_ex30	GAGGCTCTGGACACCTTCTG	GGAACCATGGGCAGACAC
NOTCH2_ex31	CAAATAGAGCTGTTTCAACCATAGG	CGATAAAACATTATAGAGCCATAGG
NOTCH2_ex32	TGAGTTTCATTATTTTCCTTCATCC	CTATTCCCCTCGTCAGAGGC
NOTCH2_ex33	CCTGCTGATCTCTGGGGAG	CGGGAATGGGCTTATAACTG
NOTCH2_ex34.1	GAGGCACAGCCTATCTGTGGTG	TACCCTTGGCATCCTTTGCC
NOTCH2_ex34.2	TGGGCCCAACAGATCTTTCC	AGTGCTGGCCCCATGTGC
NOTCH2_ex34.3	TCCCCTGGGATCTTACAGGC	AGGGCTCCCGAGGGGTG
NOTCH2_ex34.4	TTTGGTATGGTCCTGGCTCC	GCTCAGCAGCATTTGAGGAAGC
NOTCH2_ex34.5	ACGGGCAGGTAGCTCAGACC	CCGGATGACCTTCATTTGTTCC
NRAS_ex1-fw	ATGGCGAAAGAATGGAAGC	TTTCTGGCATCAGTGAAACG
NRAS_ex2-fw	GATGTGGCTCGCCAATTAAC	GAATATGGGTAAAGATGATCCGAC
NRAS_ex3-fw	GTTAGATGCTTATTTAACCTTGGC	TGTGGTAACCTCATTTCCCC
NRAS_ex4-fw	CCCAGGCTAATCTCAAACTCC	GCATAACAACAAAGAATATGAATATGG
NRAS_ex5-fw	AATGCAAGAGAGCTTATAATTTGG	TCCTCCAAATTGCCCAATAC
NRAS_ex6-fw	GCAATGACCAGGAATTTGAAG	AGTCAGGACCAGGGTGTCAG
NRAS_ex7a-fw	GCCACTTTCAAGGTAGGACAAG	TCACCAGCAGTTGCTACTTTAG
NRAS_ex7b-fw	CAGAGAACCAAACCGCAAAC	AAATGGCATCTGCTCTCAAATAG
NRAS_ex7c-fw	TCGGCTCTATTCTCTAAACTGG	GGTTTATCAGTAAGCTAGGGGAAAC
NRAS_ex7d-fw	TCTGTGACTTATCTACCATTGAAGG	ATCAAGCCCCTATTGCTGTG
NRAS_ex7e-fw	TTGGTGCATTGTTATTTAATAGCTG	TCAGCTGAGACATCCTCATTC
NRAS_ex7f-fw	TTTACCATGACTGGAAAACTCAG	TCTGGACACAGTATATGTTTCTTCTC
NRAS_ex7g-fw	CTTACTGAGGTCCTCCATCTTC	CATGCCCGGCTAATTGTAAC
NRAS_ex7h-fw	TTAGAGTTTTCCAAGTATGTTCTAAGC	TTATTTTCACTCTTGAACATGTAACC
NRAS_ex7i-fw	AGGCTATGTTTGTGATTGCTG	GGTTGTAGCTGCCAACTTAGG
NRAS_ex7j-fw	TGTCATGACTTGAAAGCAAGG	CAACTTCCACGGACATCCTC
NRAS_ex7k-fw	TGTGCCACACAATTATCTAAGG	TTCCAACACTTCCAAATGTCAC
NTNG1_ex2	TCGCAGATCATAAAGCAAGC	ATGCATATGCGATTAGACCC
NTNG1_ex3_1	CAAGACTTGTGGCAAACCAG	TGCAAATGATTTCTAAGACCG
NTNG1_ex3_2	TGGCAGCCCTATCAGTATTATG	CAGCTGTGAGACACTCTGGC
NTNG1_ex4	CCTCAGTGGTCACCAAGACC	TTGTAAGCAAAATTGAGAGGAAC
NTNG1_ex5	TGTGTGTGCTAATTTCTTAATAGTCTG	CATCAGAAATGGACACTGGG
NTNG1_ex6	TCTGCTTCTCCTCCCCG	TTTGGGTTGTGTCTGCTTCC
NTNG1-iso1-_ex6	ATGCTATTTCCGTGCCAGTC	GAAAGTTCATTCTGGCTCAACC
NTNG1-iso1-_ex7	ATAGCCATATTTGCCGAAGC	AAAACTTGTCAGCGGTTTGG
NTNG1-iso2-_ex6	TTTAAGTGCATCACAGCCAAG	TCATGATAATGCTCTTATTTTGTGC
NTNG1-iso2-_ex7	TCTTAATTTAGGGCTACTTTTCAAGTG	CATGCCTCACCAATAAGCTG
OLFML3_ex1	CTGCAGTGAGTGGAGAGCC	AGAGCTGTCTCTCGTCCCG
OLFML3_ex2	TGAAATGCTACAAACCAACACTC	TCCTTTTGAAAGAATGGTCG
OLFML3_ex3	TAGTGATTGCACAAACCTGC	TTTCTAGCTGAGTGCCCCAC
OLFML3_ex4	TCACATGACCGCTTAAATTTTC	CACTTCACACTTAGAAAATGGGC
OLFML3_ex5	CGTATTGGAATCATTGGCTG	CCCTGAAGAAAATGAATAGTCCAG
OLFML3_ex6.1	CTCACTGTCTTTCTTTTGGGG	GCTCTTCATCACCTCCAAGG
OLFML3_ex6.2	TAATGGCTGATGAAATCGGG	CGGAAGGGGTCTTATAGAGTTTATC
OVGP1_ex1	AGCGCTATCCACAGGGAAG	TGACGAACTCATAAAGCCTGG
OVGP1_ex2	TGTGCTGTGCCATACCACTC	GTGGAGGACTTCTCAGCCAG
OVGP1_ex3	GCGTTCACATTGCTTTCTGG	TGAAGAGTAACAGGAGAAATAGACTG
OVGP1_ex4	TGCCCACCGGAGTGAAG	CTAACCAGCCTGCTCCACC
OVGP1_ex5	GCCTTGTAGGAGGTTGTGATTC	GCTCTTCTCCTGCCATAAGC
OVGP1_ex6	GCTTCCTAGATGACACTCCCC	AGCTCAAGACACCTGCTTCC
OVGP1_ex7	TGCTCTGACTTGAAAATCTGG	TAAGAGCCAATGGCCTGAG
OVGP1_ex8	AGGAAAGATCTCAGGTGGTCC	CTACAGCTCTCACCCTTGCC
OVGP1_ex9	TTGGGCAAGGAACAACTGTC	ACCCCATAGAAACCGGTAGG
OVGP1_ex10	GCTACATTGCTTCCTGAGGG	CAATGTCATGTTGCCTCACC
OVGP1_ex11.1	TTGGAGCAGGTACTTCAGCC	TCAGGGTCTTCTGTCCAGTG
OVGP1_ex11.2	ATCCCTTGGAAAGCACACTG	GACCCAAGTTACCCATCCTG
OVGP1_ex11.3	CCCTAGAAGGAAGGCTGTGG	AGAAGGCTTCCAACATGTCAC

Name	Forward primer	Reverse primer
C1orf62_ex11	TTTCCCTATCTATTGGAAATGC	AAGCTGGTGCCTACTGCG
C1orf62_ex12	GCCAAACACATACCCTGCAC	TTAATTGCCTGGGAACAGTG
C1orf62_ex13	AGAAAGGCGAGTGGGGAC	TGTGAATGTCCCATACCTGG
C1orf62_ex14	GATACATAGGCACCATTTGGG	CTGTATCCACCCAACCAATG
C1orf62_ex15	AACAGGAGTGTGCTTGAGGG	GCCAGTTCTGTGGGCATC
C1orf62_ex16	GAACCCAAAGAACAGAAACAGC	GGGAAGATCAAGTTTTCTTTGC
C1orf162_ex2	GCTCTCCCTTCCAAGCTTTC	CCATTACAGCCCTTCAGCAC
C1orf162_ex3	CCTGGGGAAACTTAAGCTGG	TCAGCACAACATCAAATCCC
C1orf162_ex4	GCAGGAAAGGTTTCACTCTTTAG	GCTAATTCTAGTATCAATGGCAGAG
C1orf162_ex5	GCAGACTGAAGATTCAAATCCC	TGTCTGGAAATCAATACCACC
C1orf162_ex6	AGCAACCTTTTGCTCCAGAC	TGTGAAGATAGGGCTGAGATG
C1orf183_ex1	GGAGCGCAGTTTCCAGTG	GAGACCAGGTGCCCGTC
C1orf183_ex2.1	TAAGTTCAGGGCTGGGAGAG	ACAGCTTTGATGTGGCTGTG
C1orf183_ex2.2	GTCAACCTTCTCTTGGCAGC	GGTCAGGCCGCACACTAC
C1orf183_ex2.3	AACCCAAAGGAGAGAAAGGC	CCAGCACTGGGACCTGG
C1orf194_ex1	GGCTCGTGTGAAGAACACAG	GCTACTCGGGAGGCTAATCC
C1orf194_ex2	CAAAAATTAGCCGGGCATAG	AAACCATGATCAGGCCACAG
C1orf194_ex3-5	ACGGAGTTGGTGGATTTCAG	CGAGATGGGTAATTTCCTTCC
C1orf194_ex3*	CCCAAGAGAGCCTTCACTATAC	TGTTCATCTTAGACAGTGAAAATACAG
C1orf194_ex4*	AACTACCTTGCCTTCCTCACC	TCTTCCCTACTGTCCAATCACC
C1orf194_ex5*	CGGTTTTCCCAGCTGTTATG	ATCGTGCCATTGCACTCC
C1orf88_ex1-2	ATCACGGGACTAGCCTTCG	ACGAAACTTGGAAACTTGGC
C1orf88_ex3	GATGGCCCTGATGTTCTCTG	CTCTTGCTCTCTGCCCTTTG
C1orf88_ex4	GGTGGCTGAACCATAGCTG	ACTGCGATGGTTTGGATGTC
C1orf88_ex5	AGGTTAGGCCTAGTCCAGGG	TGAAGTTCATTCACAGACCCC
C1orf88_ex6	TAACTGGGAAGACCAGAGGC	AGCTCTGTCCTCAGTCCTGG
C1orf103_iso1-ex1	CCATTGGGAAACTCGGC	TCCACATCCCTAAGACGGAC
C1orf103_iso1-ex2.1	TCAAAATATTTGTTTTGCAGTCTTTC	GAAAAATTTCCAGTTCCCCACAG
C1orf103_iso1-ex2.2	TCAATTGCCCATTTTTCAGC	TGCTATTTCTGTAACAGGTTTTGG
C1orf103_iso1-ex2.3	CAGTGCAGCAAAAGATACTTGC	TTTTTAGCCAACAGATAGACTTTCC
C1orf103_iso1-ex2.4	AAACTTTTGTAGATAGGAAAAATTTGG	CCAACTGGAAAGATTGGATTTG
C1orf103_iso1-ex2.5	CCCAGATGGAGACAAAATCAC	GTCTGTGTGTGATGGGGTTG
C1orf103_iso1-ex3	GGCAAGGATAATACAAACCAATC	GCACAAGGGTTAAAACCAAATC
C1orf103_iso1-ex4.1	GTACCCGGTGCATTTAGCAG	GGTGGTGTCACTGGGAAAATATC
C1orf103_iso1-ex4.2	CGAGTCACAGCAAAACCAGAC	TCAAAGACACTTTCAGAACACACC
C1orf161_ex2	TCTACTCACAAGTGTTGCACTCC	TTGGAAGTAGTGTGAATGGGG
C1orf161_ex3	TTATGCTGGTATCTTGGGGC	CACACTGAGATTCACCTGCC
C1orf161_ex4	CGTTACCTCCCGAATCTGTG	TGCAGGTAACACCCAAACAG
C1orf161_ex5	AACGCTTGTAATGTGCAAACC	TGACTGGTTAAGCTGTTGGC
C1orf161_ex6	CAGAAATGGCTTTCTACACCAG	GCCACAGGAAGAATGACCTATC
C1orf161_ex7	TCCTTCCAAGGCCACATAAG	CATTCGTGGCCTTTCTCC
CAPZA1_ex1	TCCATCACTCGGCTTTCTTC	GGCCCCAGTTCCCTGTC
CAPZA1_ex2	TATGAAAAGAAAATTTGAAGACTCC	AGCTATGGCTACAACAAGTCG
CAPZA1_ex3	TGAAGACTAATTTTACAAAGCCTGAC	TGTTTTCAGGGTGCTAGTTTACTG
CAPZA1_ex4	CACTATTCTGTGAACCACCTTTTG	GGTTCAGACATCTGCTTTGG
CAPZA1_ex5	TCTGACTAAATGAAAAGTTAACTGCC	GTTTGGTAGGCCAAAGCCTC
CAPZA1_ex6	CTGCTTTTGTACAATGCATGTTG	CCAAGTCTCAACTGCTGGTTTC
CAPZA1_ex7	CAGTGGCAAATCACCTCTTTC	TTAATACCAAAGCAACCTGCAC
CAPZA1_ex8	AGCTTAAAGCATGGATTTGTGTC	CTGGCAGAATCAGCACTCAG
CAPZA1_ex9	GATGGACTTACTTTCAGGATCTTACTG	ACATTAAAGGACACAAGGCCTAAC
CAPZA1_ex10	TCCTTCACAGGGGACTCAAG	TGAAAGCCCTAGCAAAATATCAC
CASQ2_ex1	GGCTCAACAAGGCCTCTAAC	ATTCCCTCGGCACCTCAC
CASQ2_ex2	CCCTTCCATTGATACATGAGG	TTTCCTTTTGCAAGACACATTC
CASQ2_ex3	CAGCCAATACAATCACTCTTCC	CTTTGGGGTTCTATCTCTCCC
CASQ2_ex4	TGTTTCCCCAGTGAAGAAGG	TTTTGGGGTAACCTGACATACC
CASQ2_ex5	CAATCAGGATTAATTGGCTTTC	TTCAAACTTTAATTCTTCATGCC
CASQ2_ex6	TCTTTGCCGGAAAACATACAC	GTTGTTGCTCTTGGCAGGTC
CASQ2_ex7	ATAATGCTGTGGCTGCTGTG	CTTGGTTTGAGTTTGGGGAC
CASQ2_ex8	GTCCTTTGCATCCCTCCAG	AGAAAGGTGAGGGCTGGC
CASQ2_ex9	TCACACTCTGCTCTCCACATTAG	AACGCAATCATGAGGTTGTG
CASQ2_ex10	TGCTCACAGCTGTTTTGAAG	TTCTGAAAAGGCTGGGCTAC
CASQ2_ex11	TATCCCTGGTCTGGTTCAGC	TGCTGTCTGTATGGTAGTGGG
PHGDH_ex1	CGCAGCTTCTTGGCTTAGG	TACGGGGTGCATGCGAG
PHGDH_ex2	CGGCTTTACGAGTTCTCACAG	TAGGTGTGGTCAGTCTCCGC
PHGDH_ex3	GGGAAGGAGTGGGAATACTG	TTCTGGGTCCATTTGCTTG
PHGDH_ex4-5	ATGTTGCATCTCCTTCCTGG	GCCTTTCCCAAAATTGCTG
PHGDH_ex6	GCTGAGCATGGTAGTTAGTATATGG	CTCTCCCCACAGCTGGTTAC
PHGDH_ex7	AGGGAAGACCTCTGGAAGCC	GGCCTCTCTGCTCCCTATC
PHGDH_ex8	GTGGCCCAAGAGGGTGTG	AGATGCCTTCCTTCCTCTCC
PHGDH_ex9	GCAGCCAGATCTTGATTCAG	AGATCCCACACTCCCTCATC
PHGDH_ex10	GTGCCAGTGGCTTCCTG	GGGGAGAAGACTAGTGCTGAC
PHGDH_ex11	CCAAGAGAGGAGGGTGGAC	AGTTCAATTCCAGGGCAGAC
PHGDH_ex12	AAGCTTTGAACTTCTGATTCTGC	GCCCAAGAATGTGGATTTTC
PHTF1_ex1	TCACCGATTGGTTTTCCTTC	TGGCACTACTAAACTCGGGG
PHTF1_ex2	TGAGAAGCATATTTTCCACATAGC	AAGAATGCCATACTGGTAGTTAAATG
PHTF1_ex3	GGGCTACAAGAGCGAAACTG	TTGAAAATTAAGCATCCCTGTG
PHTF1_ex4	TAGTGCTACAGGTGGCCTTC	AACTGTCAAATGAAAGGGTAAATG
PHTF1_ex5	CTTGGCTAATGTGCATCTGG	CTCAGCATCAAGGAGAAAAGTG
PHTF1_ex6	AATTTGTCAGTATAATTTTCTGTAGGG	TTACTCCTATGGCCTGATGATAC
PHTF1_ex7	TTTGTCATTCTACACCCTTACTTAGC	TCCTGTGGTTTTACTTGTTAGGAC
PHTF1_ex8-9	GCAGCTTATTAGAATATTTTGGGC	AAAGGAATGGATAAAGTTATTACCCTC
PHTF1_ex10	GCTTTCAAGCCATTGCTGTAG	GGATGATTGTAAAGTTAAATGAGATG
PHTF1_ex11	ATGAATGGAGCCTGGGAAAG	AAAATAGCTCCTGATCACTTCG
PHTF1_ex12	TGTTCAGGTATGCTCTACAGATATTC	CAATCCTTTATTGCCCACATC
PHTF1_ex13	CGTTGGCAATTATATGAATACACAG	TGCCAGGAGCTTCAGTGAG
PHTF1_ex14	TTAATAATGTGAGGGCGGTTG	TTAAGGACAGTGACCTAAGTCTAAAAC
PHTF1_ex15	AGTTAATGAGTATTAAAGCTCTTCTGC	CCTGAATCCTCTGCAGGAAC
PHTF1_ex16	CATGACCTTTGAAATATGCCAG	ACCCCTCCCAACCCAAG
PHTF1_ex17	CAGCCTTAAGCATTCATCTATACC	CACGTGTGAAGAACACAGATTTC
PHTF1_ex18	TAAATCCCACCTGGTTTTCC	TCTGCAGTCATCACTTTCCTTG
PPM1J_ex1	GTAGCAGGGAAGGGCCG	GGGGAACGCGTCCTAGC
PPM1J_ex2	GACGGTGTGTGTCTAAGGTGG	AGAGGCCTGAGTCCCTGG
PPM1J_ex3	GCCCAGAGTCTTGTTATTCACC	GCTCTCAAAGGTGTTGTGGG
PPM1J_ex4	TTGATTTCTACGGGTGGAGG	CCTGCCACCTTACTCTCCTG
PPM1J_ex5	AGAGGGCCTGGTCTTTCTG	GTGCAGGGAGGCACAAG
PPM1J_ex6	TCAACGTCATCTGTGTGCC	CGAGGGACTGGTTGTGTTG
PPM1J_ex7	TAGGGGAGAGGGGCTTTG	CCTACCCTCAGGTAGAACCTCC
PPM1J_ex8	TGGAAGGGATCTTTGGTTTC	GATGAAAGCTGGCCTGAGAC
PROK1_ex1	ACAAGGCTGAGCGGGAG	TGCAGACCCTGACATCTTCC
PROK1_ex2	GATCTACCCTTCCCTTCTGC	CAGGACATCACCCTCTGCTC
PROK1_ex3	GCAGAGACTGCTGCCAGG	AGTCATGGGAGCTGGGTTTC
PRPF38B_ex1	CCGAAGAGCGAGATCGAG	CCGTTTCCTCCATACCCC
PRPF38B_ex2	GAGCTGCTCTATTATGGACTTGG	TTTTCTAACCCTATTTCACTGAACAC
PRPF38B_ex3-4	TCAATTGGTGTTAAGAGTCCG	AGTTCTTGAAGACATTAAACCACC
PRPF38B_ex5	TTGCATGCACACACTAAGCC	AAGGAAGAAAATAATCTTTGGGG
PRPF38B_ex6.1	TTATTTTCTTCCTTTCAGGGC	CCCCTTTCCTATCACGACTTC
PRPF38B_ex6.2	GCGAAGATCCCGAAGTATTG	TTTCTCTTTCTTGGAATCTCTTTTG
PRPF38B_ex6.3	GGTCCAAGGAACAGAGAAGTAGG	TTTTCTTTGCTGGGACTATGTTC
PRPF38B_ex6.4	AGAAATGCAGGGAAACGAAG	CAGGGAGAGCTTGAGGAGTG
PSMA5_ex1	CTTAGTACTGCGGCCGTGTG	CAGGTCCGGCCAGTCTC
PSMA5_ex2	TGCGTATTCTGTAGCTATCTGAGC	ACTCCAAGTGTTCTGCCAGC
PSMA5_ex3	TGTGGCCAGTTGAAGTTTCTATC	AGCTAACAAGGGGATAACGTG
PSMA5_ex4	CCTGAGTGGATGAAAGAGTTTTAC	CCCAGGCAGCTAGATAATCAC
PSMA5_ex5-6	TGGTGGGCTGCTTTAATGTC	AATCTTCTCAGTCTCCACCCC
PSMA5_ex7	GCAGGGCCAAGCATTTATAC	GGAGGGAAAGAGATTGACTGG
PSMA5_ex8	TTTGAAGGAAGGACGGATTG	GCAATATAGCCAGACCTCATCTC
PSMA5_ex9	GGCCATCAGCAGGGTAACTC	CCAAGGAACAGGAGCTGG
PSRC1_isoA_ex2-3	CACCCTGAATCTAGCTGCG	CCAACCCAAGCCAATAATTC
PSRC1_isoA_ex4.1	CTTGCATAATCTCCTGTTCCC	CTTCAGCACAAAGGTCTCCC
PSRC1_isoA_ex4.2	AGATCCTCGATGAGGCCAAC	ATCATGGCACACACCATGTC
PSRC1_isoA_ex5	ATTTCTTAATTCCTTAAGCCACAG	ATCCCACAAACTCCCCACTC
PSRC1_isoA_ex6	AGTGGGAGTGGGGAGTTTG	TCACTGTTGACAGCTCTGGG
PSRC1_isoA_ex7	TCATCCAGTATTTCTTTTCGAGG	CTTGGAATCCCAGATGATGG
PSRC1_isoA_ex8	GACAAATGGGCTTGAGGAAG	AGGGTCTGCTGGGTAGAGG
PSRC1_isoB_ex8	GACAAATGGGCTTGAGGAAG	AGAATCTGCTGGAGTCAGGG
PSRC1_isoC_ex5.1	ATTTCTTAATTCCTTAAGCCACAG	GAGGGAATGGGCAGTTGAC
PSRC1_isoC_ex5.2	ACCCTGAAATTTCTCCACCC	TTGACAGCTCTGGGAGGG

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
CD2_ex1	ACGTGGTTAAGCTCTCGGG	AATCTCTTTGGAGACTGCACC			
CD2_ex2	TCAAAGAAGTCCCAACCCAG	CCAACTTGCCAAATAGCACC	PTGFRN_ex1	GTTCTTCCCTTCTGCTTTCG	AACAGGTAGGGAGACCTGGG
CD2_ex3	CCCTAACTGACTCCATCCCC	AGCATGAGCACCTGTGACC	PTGFRN_ex2	CAGGTTTGCCACATTTGTCC	AGGGGCAGAGAAAGGTCTG
CD2_ex4	GGCCAGAGTAATGGGCTCTC	GTGTCATCCCCATGAGCTG	PTGFRN_ex3.1	GAGGAAATTGTCTGTGGTTGC	ACACTGAGAGGCGGTAGGC
CD2_ex5	GCAACTTCTGCTTCCTCATTG	TGAGTTCGAAGATGCAGAGG	PTGFRN_ex3.2	AGCGTCCTCGCCCTGAC	GGCCCGGTTCTTTGTATCAC
			PTGFRN_ex4.1	AGGGTGGAGAATAGAAGGGG	AAATGGTAGGAGCGTGCATC
CD53_ex3	TGCTAGAGATCCCTGAACATTTG	CATGCATTTTACCCCACACC	PTGFRN_ex4.2	GATGCCTGACAGCACCCTAC	AACTCCGAGGAGGGGTTAAG
CD53_ex4	TGATTCTGATCTCAAGGAAGTGAG	TGGGATTTAGGCACCATCTC	PTGFRN_ex5.1	TTTAATTGTTAAAACAAGAGCAGTCC	TTTGGATCCGGAAATTGAAC
CD53_ex5	GAGTGGGAGTGGGACGAG	TGCATACTCCCCACTTTCAG	PTGFRN_ex5.2	GACAATGTGGTGACCAGCG	CGGCACGACTTAAGCAAAG
CD53_ex6	TCCTTATTCCTTGAACTCACCTG	CAAGGATTTAGGCTTGAAAATG	PTGFRN_ex6.1	TCGTAGGTGCTTTTGTGTGG	TTCGATAGCGGAACTCATCC
CD53_ex7	TCTGAAGGGCACACCTTTTC	CTATGACTGCAGTTTCCGAGG	PTGFRN_ex6.2	CCAGTCCCAATGAAACGAAG	TAAAACTGGTCGTGGTGATG
CD53_ex8	AAGCAAAATGCAAAGCACTG	ATTTGCTACTTGGCTGGGG	PTGFRN_ex7	AGAACCACATTTTGATGCCC	ATAGATAATCCTATCCCTAACCCAC
CD53_ex9	AATCCCAGGCATTGTGAAAG	TCCAGTTTGTCTTTGCCTCTC	PTGFRN_ex8	GCCATACATATGACTTCTGCCC	TTCTGTCAGCTACAGCCTGG
CD53_ex3*	CAGCTCAAACCCAAGGTAATG	GCTGGACATGACCCAAACAC	PTGFRN_ex9	CTCTCGGGAGATCCTGCTC	CAAGTCCCACCTGAGATTAGAG
CD58_ex1	CGTAGGCGGTGCTTGAAC	GTGCTGCCCAGTACCCG	PTPN22_ex1	ACATTTGACATGCCCTCCC	TTGGTCTCGGAAAATTGGAC
CD58_ex2	TTGTGTCAGCAGTTTGTCAGC	GACAACAGGTAACATCTAATTCTTTTC	PTPN22_ex2	CCAGTGGAAAAGGAAAGCTC	GCCATGTTCCAAGATCAGG
CD58_ex3	GCTTGTAAGGAGGTTGCGAG	TTGTGAACCTTGTGTTAGTCACC	PTPN22_ex3	TGGAACATGGCATTGACAAC	GTTCTCACCTAGTCCTCCGC
CD58_ex4	GCTGGGAATCTGCTTATTGG	TGTGGATCTTTGGAGGAACC	PTPN22_ex4	CGGAGGACTAGGTGAGAACAG	CCTGAACCCTGAAGTTCCAC
CD58_ex5	GTGATGGGGCTCTGGCTTAG	TAATGGGCATCTACAGCAGG	PTPN22_ex5	GGATCCATGTATTAACTGCTCTTTC	CTGCCAATTTCCCTCTCC
CD58_ex6	GGTCTGCAGCCCAGATTTAC	CAGCTGCTTCAAGTTACATTTCC	PTPN22_ex6-7	ACATAGAGCTGAATTTGCTTCC	CACTTCCCATTGTCTTATGCC
			PTPN22_ex8	GCATGCCACTGAATAATCCC	TGGCCAGACTCTGACATTTG
CELSR2_ex1.1	AGGTGGGCGATCCCATAG	AGTGTGAGCTTGCCCTGTG	PTPN22_ex9	TGTTGCCTAGGACCTCTTGC	ACCACCCTTTAAGACCCACC
CELSR2_ex1.2	CGAACCTCTGGCTCTACACC	TTGGAGCGGCTATCAAAGAG	PTPN22_ex10	CAGATGCCATGGAAGTACAAG	TTGGATGCCTCACCATTAAAC
CELSR2_ex1.3	ACCAGAAGAGTCCCTGGGTG	CTCGAACACAGGGTCATGG	PTPN22_ex11	CGATCGTCCTTTTGGGTC	TGCAACAAATCTGACATCCC
CELSR2_ex1.4	ACCCACGTCTTCAGGGTCAC	CAGCGGCTGTGGTACTCC	PTPN22_ex12	ATTTTGGAAGCTAATCTAACTCTTTC	GACACCATACCCCAAACAGG
CELSR2_ex1.5	CAGCCCCTCTGAAGTCTTTG	TCCAGGACCTGTACTGTCACC	PTPN22_ex13.1	TAGGAAGCGGCCTGTTTG	CCGTGTTATTGGCACCTTTG
CELSR2_ex1.6	ATGTGGTGAGCCCTCTTGAC	TCTGGTTGGGTAAAGGTTGG	PTPN22_ex13.2	TGACACAACCATGAAATGGC	ACTGGTACCACTTGGAGGC
CELSR2_ex1.7	TCTCTGTGGCTGCTGAACTG	GCCGGTCCTCATTAACATTC	PTPN22_ex13.3	AAGTAATGCATGTTTCTTCAGCAG	AGCAGAGGAGAAGAGGGAAG
CELSR2_ex1.8	CCTGCCACTGGACTACAAAC	TCTGCAGGACGCTAGTGAAG	PTPN22_ex14	AAAACCTCCTGGGTTTGTACC	CATGCTGCTATTGCTCTGC
CELSR2_ex1.9	TGAACGACGTGAATGACAATG	GCTGAAAGACCTCAGGGATG	PTPN22_ex15	AACTAGATTACTTCTCAGGTCCTTTC	AGGATTTATTGAATGATGGGTG
CELSR2_ex1.10	GAGCAGGATGAGTTTGATGTG	AGTGCGCGGCTTAGCTTC	PTPN22_ex16	TCAGAAGGGGATCAATTTAGG	TGGGAACCAAGTAACTTCCTG
CELSR2_ex1.11	CTTGACCGCAATGACAACC	GCAGGGAGGTGGATCTGTAG	PTPN22_ex17	ACAAAGAGAAAAGGGTGTTGAC	TTGTACCTTTCCATTTAGGTCC
CELSR2_ex2.1	ATGCCAGTTGGAAAGGAGAG	GCAGGCAGATGTTGTCGTC	PTPN22_ex18	TGTAACATGAAGGCTTAAGAAGG	GGGGAAACTTTCAGTAAGGG
CELSR2_ex2.2	ATCCTCAACGTGAGCCTGTC	TTTGTCCTGATTTGTGCCTG	PTPN22_ex19	AAAACAACTTTGTACATTTTCATGG	TGGGTGATTGTTCTGATTCATAG
CELSR2_ex3	GGGTCTCAGTTCCCCTCC	AACCACAGAACAGGAGGCAC	PTPN22_ex20	TCAACCCATTTGTAAGCACC	TGTGAGATGTTAAGAGGTAGTTAAGG
CELSR2_ex4	GTGCTGGCATCCAGGTG	AGTATGTGGCTTGGGGTCAG	PTPN22_ex21	AAACTGGTCAAGATGCTGCC	ACTTATTGGCATTTTGCTTTTC
CELSR2_ex5	AAACCCTGCCATTCCTAGC	ACTCAGGGCTGGGTAAGCAC			
CELSR2_ex6	GCTGGCCATGAACCTAGTG	CTCTCCACCTTCCCTCCC	RAP1A_ex2	ACTGTTAGCATGTTTCTTAATCTTTG	ATGCCACCTTTTCTCTTTGG
CELSR2_ex7	GAGGACCAGAAGCTCCCAG	AGCCCCTCCTCTCAGCAC	RAP1A_ex3	CGTGTTTACCCAGAATTTTCAC	TGTTAAGAAATGTTCTGTCTTGATG
CELSR2_ex8	GCCTCTCTAACTGCTGCCAC	GCCAGTTCCCAGACTGTTG	RAP1A_ex4-	TTGATGGAAGAGGTGGGAAG	AGGAAGAAACCCAACACACAG
CELSR2_ex9	CCACCGCTGAGCATCAC	AGCCGCCTAGAGGCTACC	RAP1A_ex5	ATGCTGTTTTAAATTGTTGCAG	TGTGGTAGATAGCTTCCAAATCC
CELSR2_ex10	GCTGGAGAACAGTGATTAAGGC	AGCAGTGTTGCCCTCAGTG	RAP1A_ex6	ATTTGATGAAGCTTGCGGTC	CCATCAAAGTGACCTTATGCAAG
CELSR2_ex11	GGTGAGTGTCCTGCCCTG	TCAGAGAGGTGGTGTGATGC	RAP1A_ex7	GAATTGTTTTGTGGAACAAAGTCTC	AGGAGCATGAGAAAGGCAC
CELSR2_ex12	GGGTTTGATGGAAGATGGTG	ACACCAGAGAGAAGCAGCAG			
CELSR2_ex13	CTAGGTTAGGTGGGAGTGGC	CAGGAGATGTGGACAGAGGC	RBM15_ex1.1	TGGGAACCCAATAACAGTGG	GCTCGGGGAGCTATAACTATG
CELSR2_ex14	AGAAGGACCCATGAGCATC	CTGGCTTTCCTGGGGTG	RBM15_ex1.2	GAGCAGCAGCGGAAAGAC	TTCAGAGGCCGGTCATAGAG
CELSR2_ex15	GCAGGACAGCAGAGGGAG	ATCCCACTAACAGCAGAGGC	RBM15_ex1.3	TCATGAGTTCAAACGCTTCG	CCTTGGCTCAAGACTGTATGC
CELSR2_ex16	GCAGAACCTCTCAGCCAGTC	GGGCAAGTGAGACTCAGGAC	RBM15_ex1.4	GAGATCACTTTCCCCTGGTG	TGGGTGTAGCTTTACCATAACC
CELSR2_ex17	CTGAGGTCCTTTTGCTCCAG	CAGGCTCACTCACCATGACC	RBM15_ex1.5	GCGGCCAGACTAGTACTTACG	GGTGTCCTATCCCGAGCAC
CELSR2_ex18	CTCCATGGTGGGGATGTC	AGCACAGCCGCTCTTGTC	RBM15_ex1.6	ACTTGGTGGCCCAGATCG	CTACTGCTCAATTCTGGACTGC
CELSR2_ex19	GAGGGAGTCCCCGACAAG	GGAGTCTGCACAGACACTGG	RBM15_ex1.7	ATCTGCCCAGCAGCAGAG	GAGTTTGGGAGAGGCTGATG
CELSR2_ex20	CTGGACCCCAGTGTCTGTG	CTCCATGCCTCAGCACTCAC	RBM15_ex1.8	AGGGAAAAAGCCCTCTGAAA	GCTGCCTGCTTTTGCTTTA
CELSR2_ex21	AGTGCTGAGGCATGGAGG	GGGCCTGGACTTTCAGC	RBM15_ex1.9	CCCAATGGTTATGCCATTCT	TGCTGTATCCCATTACACATCA
CELSR2_ex22-23	AGACAGAGGAGCGGACCAAG	GTGTTGGGAACTGCTTCCAG	RBM15_ex2	GGGCCTTAATGACCTTTTATGTG	ACTCTTCATTCAGCTTAACATTTCAC
CELSR2_ex24	GAGCCTGTGCTCTGGGC	GCCTGTATGTGGGTGGAGAG			
CELSR2_ex25	GCAGCCCTACTTCCCAGG	AGGAGGACAGCAAAGGAGG	REG4_ex2	CCCAGGAGTTTCAAAGAAGC	CTCACATGGGCTGTGTGAAC
CELSR2_ex26-27	CTGGGTGGAAGCTGTTTGTC	ATATTCAGGCATCCACAGCC	REG4_ex3	TGCAAACTGTATCAAGTCTGGG	ATCTGGGGCTCTGGCTG
CELSR2_ex28	CACCTGGGCTGTGGATG	GGCCAAAAGTACAGGCTCAC	REG4_ex4	CTGCGTGAGTCCCTCCTTC	CCCGCTGTGTCCTTGTTATAG
CELSR2_ex29	ACTTAGGGCAAGTTCCCTCC	GTGTGTGCTCCTCAACTCCC	REG4_ex5	TTAAAGGGGACTGAGGGGAC	GGTAAAGGGAAGAGTCTGCG
CELSR2_ex30	GGAGAGGAGAGACTGGGACC	CCACCCACCTTCAGTAAAAGC	REG4_ex6	CCCAGATAACTCTGGCAAGC	CCAGGCTAGCAGAAAGGAAG
CELSR2_ex31	TTACTGAAGGTGGGTGGAGG	CAGTCCTGCTTCCCAAGAAC			
CELSR2_ex32	GGTTCTTGGGAAGCAGGAC	TGTGAGGTATGAGGCTGCTG	RHOC_ex2	TAGATGCCTCGGTTTCCAG	AGGTCATTCAGTGACTGGGG
CELSR2_ex33	TCAGCAGCCTCATACCTCAC	AAATGCCATCCAGAGTCACC	RHOC_ex3	CGAAGGACAGTGGTTTCAGG	ACGAGACCTCTGGCTGATTC
CELSR2_ex34	CCTCTGTTTCAACTAACCCTCTG	AAGCACAAGACAGGAGCAGG	RHOC_ex4	CCCAGGGGAGCTTTCTAGC	AAGTTCCCTTTGCCCGTC
			RHOC_ex5	GGTGAGGCAGGAGAGGTTC	GGGGATAATTTCTGTACCCCTG
CEPT1_ex2	TTTAGGTAAGCACCAGCCAC	GACCCCACGATCCCAATTAC			
CEPT1_ex3	GCACTTGGATCTTCTCTCCC	CCAAAACAATATTTGGGGC	RSBN1_ex1.1	TAGAAGCTCGGTTGGAGAGG	CCCCACTGCTAGATCGGC
CEPT1_ex4	TTTCCAGCTCAGTTGTTTGC	TCAAATACTTATAAGGCAACATTCTC	RSBN1_ex1.2	GTAGTACGGGCGGTGGC	AGGCGAGAAGGTGAAGAGG
CEPT1_ex5	TGATGCCTCCTTTTCCCC	GAATTTCACACATTACCCTTTAACTG	RSBN1_ex1.3	TGTTGAGCCTCTTCTCCTGC	CTAAAGATGCGGGATATCGG
CEPT1_ex6	TGTAATTGGAGGGAATGGAAAG	AAATGACAACAAAACCAAGTCC	RSBN1_ex2.1	CAAGTACCTTAAGAAGCTTGTATGTTG	CCCGATATTCTTGCATACCC
CEPT1_ex7	CTGCAGCAGATCCATGTAGTATC	GAACTATAGCCATAAATCCAAAATAAG	RSBN1_ex2.2	AAAGAAATTCTCACCCAAGGAC	TCATTTTCACTGAATGTCAAAGC
CEPT1_ex8	TGGGAATTCAACAGCTATTCC	GAAAACAAGCAACCTTCAAACC	RSBN1_ex2.3	TGGTGGTGCTCTTGTCCTTC	ATGTGCTGTGCGACAGAAAG
CEPT1_ex9	TGTGAAGGAAAGATTGCAAGG	GGGCGTGTTACCAATAATAAAG	RSBN1_ex3	GAAATGGCAACGTTGGAAAC	GGCAATATTAGCTAAATACTCTTCAAC
			RSBN1_ex4	AACTTTTATTTGGGTGGGAGG	GCCCAAGACAAATGGTTTTC
			RSBN1_ex4	AACTTTTATTTGGGTGGGAGG	GCCCAAGACAAATGGTTTTC
CHI3L2_isoA_ex1	AGGCTGTCGAAACCTCAGTG	CCTGTTACCTCTTCCTTCCTCAC	RSBN1_ex5	TCATTGAACCCTTCTCAGTCC	TCTGTTTCGTAACAATGTAAAGATG
CHI3L2_isoA_ex2	GCCATTATCTGGGACAGCAG	AAATGAAAGCACCCAAGCC	RSBN1_ex6	TCTGCTGGTAGAAGCATTTCC	AAAATTAAACTCACTATCCCACATATC
CHI3L2_isoA_ex3	TGGACTCTAAGGCAACAGCC	CAGCTTAAGACCAGCTTCCC	RSBN1_ex7.1	TTTGTTTCTACTGTTTTCCTTTGC	TGCATGCTTCTTCAGATTCAG
CHI3L2_isoA_ex4	AACAGCCAGGCCCTAATTG	ACCTGGCCAAGGGAAAAC	RSBN1_ex7.2	GGAAGCCACTCAAAACACAG	TTATAGAATTATAGGCAAGATTTCTCC
CHI3L2_isoA_ex5	GGGAAGAGAAGACTGCTCAGG	AATGTGGCCATGGATTCTG			
CHI3L2_isoA_ex6	CCAGGCAAGAAGCTTAGCAC	GATGCATCATTTGCATGGAC	SIKE_ex1	TCCCAGAAGATCTGAGCGAG	CCGAGTTCCCACTTTACCTC
CHI3L2_isoA_ex7	CAATTGTTTCTACCACTGCTCG	AGACATGCTGGACTCTCCCC	SIKE_ex2	TGAGACCATTTTCCTGGGTC	TCAGGAGGACTCTCCTACTAAAAG
CHI3L2_isoA_ex8	AAACCAGGCAGTCACCTCTG	TCTGCTGTCAAATTGAAGATGG	SIKE_ex3	TGGATTTGGGAACTGTCTCTC	CCAGACCCTATCTCACAAAGC
CHI3L2_isoA_ex9	CAAGTCCCTAGTGGCTCACC	TTGAGCTATAAAGGCTGCTGG	SIKE_ex4	TTGTAGGGAAACTCCATTGTAAC	GGCATATCTTCCAGACTCGG
CHI3L2_isoA_ex10	TATGGTCCTGGGCATTTAGG	CCCCAGTCTCCTGTGCC	SIKE_ex5	TGTGTTGTTGTAATGCAGTTGG	TGGACAGTACTTGAATGGGAAG

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
CHIA_ex2	ACACATCTGGGTCAAATTGTTC	TGAAAGAAGAATGGAAAAGTATAGGG	SLC16A1_ex2	TTTCACAAAAGAGTTTTATAGGTGTG	TTACCTAATACAGACACTCTGGCTGC
CHIA_ex3	AAATTGGAAGGCAATCAATCC	TGTCTGTGAAGCAAACAGCG	SLC16A1_ex3	CTTGTGTAGATGTGAGGGAGC	CAGGTAAATACAAAATAGCCAGCC
CHIA_ex4	CCACACAGAGATCAGGCATC	TGCAGAGTCTTGGTCCCC	SLC16A1_ex4.1	TCCTATTCCTTCCATTATTTATACCTC	CCAGGAACTGATTAATTGTTTGG
CHIA_ex5	GGAAAATTGTTCTACAAAAGGC	TTGTGAAGTCTAAAGGGACAGTG	SLC16A1_ex4.2	TTCTAATTCTTGGGGGCTTG	ACACTCCATTTGCAACAACG
CHIA_ex6	TACTCCAAGGTGTTGGGACC	ACCCTAGAGCCACCCAATTC	SLC16A1_ex4.3	GGGAAGAGTCAGCATTATTCTAGTG	AGCATCAAGGGTACATAAATGAG
CHIA_ex7	CCGATCTACTTTATTTGGTTGTTG	CCTGAATAGCAAGCCAGACC	SLC16A1_ex5	GGATTGACATCTTTCTGAATGG	ATTTGCATTGAGCACCACTG
CHIA_ex8	CCTGTGCCTGCATAGGTGT	GATGCCTTTTGTGCCTGAAT			
CHIA_ex9	ATAACTAAGTACTGGGTCCTCAGC	CAGATTCTAGCCCTAAGGCAC	SLC16A4_ex2	TCCACTGAATCAACTGTCTTGG	TAGGCAGCTCTGAGCCTTTC
CHIA_ex10	GCAAAACCCCAACTGGAGAC	TGCTTAAGCCATTGAGCCTG	SLC16A4_ex3	TTGCCTCTGTAATGGCTCAC	AATCTTGGAATTGCCCTTTG
CHIA_ex11	TCCCTTAAATGTGCTGAGGTC	GTGGGAAGACATCAGGGTTG	SLC16A4_ex4	CAGAGGATAAGGGATGCCAG	TCTTCTATCTTCTATGGATGGATGC
CHIA_ex12	TGAAGCTTAGAGCCTCTCCC	TAGGTAGGGGCAACATGTCC	SLC16A4_ex5	CGGATCCAGGAGATTATCTTTTC	GAAAGCTGAACGCTAAATCTTG
			SLC16A4_ex6.1	CAGTGGGGCCTCTTCTGAG	AGTTGTTTGCAGCTCCACG
CLCC1_ex2	GCAATAGAGCAAGACCCTGTC	TGTTTCCATCATGGGACTG	SLC16A4_ex6.2	CAGTACTACGCAGAAGGCTGG	TCTGGCCCTTCTACAATTCC
CLCC1_ex3	TCTTAGTCTTCATTAACCAAAGGC	AAAATCATACGGTACACGGAAAC	SLC16A4_ex7	TGGGAAATGATGCTGTATGC	GGGGAGATTGTTACTGATACTTGC
CLCC1_ex4	GATTCCGATGACAGCACTAGG	AAGAGTCAGTCAGGTGGTCTCTG	SLC16A4_ex8	TTGTCCCTTTTAGGCATAGGTC	CCTGATACCATCATTACATGTTACC
CLCC1_ex5	AGTTCCAAATTGGCTTTTGTC	AACGTAATTATGGTGGAGTTTGG	SLC16A4_ex9	CATGGCTGAGAGTTACTAATGGC	TTCAAGCTTTTGTTTCCAATG
CLCC1_ex6	TTTGTGCTTTTGAGTCCAGTTC	TCCAACCCACATCCAAGG			
CLCC1_ex7	TGTGTAGGCCTAACCTTCATAGG	TTCAGTCATGGAAAGCCTCC	SLC22A15_ex1	GCCGCCAGCGCTTCCATCC	CCCGACGCTCGCCACACCTC
CLCC1_ex8	GGTCAGGGAGGGTCAAAAC	CAAACGCTTATTAGCAAAATGAC	SLC22A15_ex2	TCCTTTGTTTCTGATGTGGC	GAGGAAAAGGAATGTGTTCTCATAG
CLCC1_ex9	TCAATGAATTTTGATGCATTCTG	CGATCTCCTGACCTCGTG	SLC22A15_ex3	CTGGTTAATAGTGGATCATGCAG	GTGTCCATATGCCAATTCCC
CLCC1_ex10.1	TGGCCTTGTATTAATCTAAAGAGGTC	CACTTCAGGGCTCTTGTTGC	SLC22A15_ex4	CAGCGAGCCGAGATTGTG	TTGTTCAGGGTCATACAGAACC
CLCC1_ex10.2	TAGACCTGATGGTGGAGCAG	AAAACAATCAGACCTCTTCCAG	SLC22A15_ex5	TGCCTCTTATAGTTGCCAGC	CAAGGGATACCTGCATTCTG
CLCC1_ex11	TGAAAGAGTTTGGAGCTTGC	TGAATGTTGAAGTTGGCACC	SLC22A15_ex6	TGATTTGTCTCCTTCCTGTGC	GAGAAGCTCGAGCAACCTTAG
			SLC22A15_ex7	GCAGATCAAGCAATTATATTGAACC	TGATACTGCATACAAAGACAACACC
CSDE1_ex2	GACTGTGAAATGTGCATTAAATTG	GGCAATCCTAGGTCATGCTC	SLC22A15_ex8	AAGAAGGTTGGGAAGAAACAG	TGCAATCACAGAAATGAAGC
CSDE1_ex3	TGCAGTAATAAAGTCTATAATCTTCCC	TGCTGGTTTATACAGATATTCTTCAC	SLC22A15_ex9	CCATCAAAATGAAATGTGCC	TGCAAGAATGGATGACAAAG
CSDE1_ex4	TGAAAACCACTGGTCTAAGGC	CAAACTGAGGAGATGGGGAG	SLC22A15_ex10	GGGTCTCTTAATTCCTCCGC	GCATAAGCATTTCCAAAGCAG
CSDE1_ex5	ACAGTGCTGCTCCTGAAATG	CGCTTATAGATCACATGGTGG	SLC22A15_ex11	TGGCTTATAGTGTTGTCTAGTGCG	CCACTTTTGTTTTCCAGCAG
CSDE1_ex6	ATGAATAAAGTGACTTGTCATAAAGC	TTCAAATTTCTTAATGGAGGGG	SLC22A15_ex12	AGTTGAATAGCTAGATGGTAGTTTCC	CCTTAGCCAGAGGCATAAAAG
CSDE1_ex7	TTCTTAAAGCCCCAGGAAAC	GCAACCAAGGAATTTTAAGAGAC			
CSDE1_ex8	TTCTTTAAAATTGGCTGTTTAATTG	AAATCAGAAGGGGTTGGAGG	SLC25A24_ex1	GCAGCCCTCGATCTCCC	CTCGGTCCTCGCAGTGTC
CSDE1_ex9	GGCCTTTTAAGATCAGAGTTGG	GAGGTCCTTAGTTGTTCCACTACC	SLC25A24_ex2	GAAATAATTTAGTGCATCAGCCC	ATTTACAACAGGCCCTGCC
CSDE1_ex10	TGTTTGACTTAGAACTAGGAAGGATG	TGAATATAACATTGGCTTTCGC	SLC25A24_ex3	GCAAGTCTATAACTCACCTACATGC	GAGGTGTTTTAAAGCCACACTTAC
CSDE1_ex11	GGTTGTAAAGTTGAATGTTGCC	AATCCTAATAGCATGATTTCTTTGAG	SLC25A24_ex4	CTGTAAAGTGGCCAATAATGTTAG	TTGGGTGATAACTGTACCTATGG
CSDE1_ex12	TGGGAGGTGTTTCTTGATGC	GCTGTTAAAGCAGAAGAGAAATAAAC	SLC25A24_ex5	TCTAAAACAAAACGTCTTCTGGG	CAAAGCACCACATTTCTCATTG
CSDE1_ex13	GCTGTACCAATCCTTTGGC	ATTTGAAGAACCCATCCTGC	SLC25A24_ex6	TGGCAATACTCCCTCTCATC	TTAAAACAAGAATCCCAAAAGC
CSDE1_ex14	TGAGTTAAGGAGATTGGTGTGG	CAACACCAAAGAATGGATGG	SLC25A24_ex7	CCATCACATGGGAGGTTACAG	GAGCCAGAGGCACCCAG
CSDE1_ex15	AAAATAAAGACTGCCTGTTACTGTTG	TCTACTACTCCATCCTTTATCTCCC	SLC25A24_ex8	ATTTGAATCCCATGATGATAGAG	TGGTCACACACTGTTTTCTGG
CSDE1_ex16	CCAACCTACCCTGTTAACATTTG	AGGTGCTCATCTGAAGTTATGG	SLC25A24_ex9	AGCTGTGAAAATGTTTACGTTG	AAAAGCACATTACTGGTGCAG
CSDE1_ex17	AAACTTCCTAGGGTCATCAGC	AAACGAATGAATGTTGATGATTTC	SLC25A24_ex10	GGCTTCTTGGAATAAATGTTGTC	AAATGCAGCTTCTTTTGCC
CSDE1_ex18	TTAGATAGGGGCAGGCATTG	GACTATAAAGTTGAGACACTCCAGG	SLC25A24_iso2_ex1	GTCGGTCGCACTTCCAC	AATGCCTCTCCAGGGCTC
CSDE1_ex19	CAAGTGCTTGCTGTCTGCC	TGACCTATTGTTTAAATAGGACCAC			
CSDE1_ex20	GTGCTTACTGTGGATTTTCTCC	TCGGGAGGGATGAAGAGG	SLC6A17_ex2	CCCTGAATGAGAAGGAGCTG	CATCCCCGGCCTATTTG
			SLC6A17_ex3	TTCCCTCACTCTGCTGCTG	AAACTCGGTGACCTCGTCTC
CSF1_ex1	GAAAGTTTGCCTGGGTCCTC	GCAAAGGCGGTTCGTCC	SLC6A17_ex4	GGGGCAGGGTGGAGTTG	CGCTAAGCTTGTGAGAACCC
CSF1_ex2	CATGGGGATAACTGGGGAG	CTCTTTGACCTGCTCCCTG	SLC6A17_ex5	GAGGAAGTGACCCCATAGGC	ATTCTTAGCCCTGTGCCCTC
CSF1_ex3	GAAGGCTGAGTTGAGCTGTAGG	CCGCTTTCATGCTCCCC	SLC6A17_ex6	GTGCCAGCTGCAGGAAG	GAAATTGCAGGAAGCTGACC
CSF1_ex4	CCATCTGCCTGGGGTTG	CTCCCAGCCACCCTCTG	SLC6A17_ex7	AAAACCATCCCCTCTGTGTC	CCCACTCCCAGGTAAGGTAG
CSF1_ex5	AATAAGATGGAGACATGGGGC	CACATCTCACCTCCCAATCC	SLC6A17_ex8	CTCCAGGATCTCACCCATTG	CACAGCTGCTCTCTGGGTTC
CSF1_ex6.1	GCTAGTGACTCTATCTCCTCCCC	TGGGGTACGGGAATCTCAC	SLC6A17_ex9	AGAGGGAAGGAAGCAGTGG	GGAATTGTTGTCCACAAGGG
CSF1_ex6.2	ACCCCAGTTGTCAAGGACAG	GAGGAGTGGCTTCTGGAAAG	SLC6A17_ex10	TTGGGGAGGGATGTGATG	TTCAGCTCTGCCATTGTGAC
CSF1_ex6.3	ACAGAGCCCGCCAGACC	GTCCTCCTTCTGGCTCTGC	SLC6A17_ex11	GGAAGAGGGAGGTGTGACTG	TGTATCCCTATCAGTGGTTGC
CSF1_ex6.4	CAGCCACAGCTTTCCAGAAG	GAAGAACTCAGGCCGCAC	SLC6A17_ex12	CAGGAGAGCAGTCTTGGAGG	GACAAGTGGGCAGGTTGAG
CSF1_ex7	ACCTCACAAATCTCCCTTGG	CAGGAGCAGGGGAATAAGG			
CSF1_ex8	CCCTTATTCCCCTGCTCC	CAACAGGTGAACGTGACTGG	SORT1_ex1	TTAGCATCCGAATCCAGGAC	AAAAGAATGCGGTGGAAGTC
			SORT1_ex2	GAAGCCAAAGATTGAAGGAGAG	TAGGTGCTTAACCCTTCCCC
CTTNBP2NL_ex3	TGTTGTTGCTTAGTTTGTGATTAAC	AGGTGGCTACCTATGGCAAG	SORT1_ex3	CATATCAGCATGTGATTGGAATATC	GCCATCATGTCTAGAAAAGGG
CTTNBP2NL_ex4	ACCACTGGATAGCAAATCCC	TGGCTCTTACATAAACTAAGGGG	SORT1_ex4	GGGAGAATGAAAGCTGAGTAGG	CTGACAGCTCTGCATGTGG
CTTNBP2NL_ex5	TTGGTTGTATTAATAAATGTGACTCC	TTTGGGAAATGCTGTTCCAG	SORT1_ex5	GGTGTAAGAACTCCTGTGCTTG	GAAGTCCAATATTAAAAGCATTTCC
CTTNBP2NL_ex6.1	CCAGAGAACAGCTAACGAAGAC	TTTGGTCCGGTTCTCTTCTC	SORT1_ex6	GGTCATACAAGCAACTACACAGG	GCAGCGCTATCAGGAAGG
CTTNBP2NL_ex6.2	AAGAGAGGCTTGCAGACTGAG	AGTCAGCTCCCTGGTAGTTTG	SORT1_ex7	TCAACACAGTCCTTCAGCAGTC	CAGAAAATTCTACCATGAATGGAC
CTTNBP2NL_ex6.3	CAAGATGACAAACACTGGGC	AACCGTTGGTTGATCCCTAC	SORT1_ex8	CTGCTTTGAACTAGGCTCACC	GAGAATCTATACGAGAGATATTTGCAG
CTTNBP2NL_ex6.4	CATCCTCTCCTTGCTCTTCG	TCAGGGGACTGGACACTTTC	SORT1_ex9	TCTTGTTCATGTCCCTGCC	AAAATAAAAGCTGCTCAAAGCTTAC
CTTNBP2NL_ex6.5	CCCAACAGCTCTCCCTTTG	TCTGACTGAGCTTTTGGACG	SORT1_ex10	ACTGTGTTTGGTCAGGCAGG	CCCGAAGCCTCAAATCCTAC
			SORT1_ex11	CCAGGTGCTATACGGTCTCAG	TTCATCATCCAAAAGGCTGG
CYB561D1_ex4.1	AATGTTGAGCTGTGAGGCTG	ACAAAGGAGGCAGAGCCC	SORT1_ex12	AAGGGTTTTGTGAAATGATTAAAG	TGGCTGAGTTTCAGTGTGC
CYB561D1_ex4.2	ATCATCTCCAGCAGGACCC	CAAAGGAACCATGATGTTCAAG	SORT1_ex13	TTCATTCAACATTATTCTCTCTGG	AATTTGAATTATAGTCTGTAACATCCC
			SORT1_ex14	GAGTGGTTTTCTGAATTCCCAG	TGATTATTTAGGAAAATGGACTTGG
DCLRE1B_ex1	GTTGTGGAAGCCTCACGC	AGGTTGAAGTCCCAACCCAG	SORT1_ex15	TCTTCCTATAGCCTGGCAGC	CTTTCACCTCCTCCAACAGG
DCLRE1B_ex2	GGATTTGGTCAGCTTTCCTG	ACTGCGAGGCCCCAGTC	SORT1_ex16	TCACTTAACGGTCTTCCAGC	GGATCTCAGTGTGGAGAGAAGC
DCLRE1B_ex3	CGCCTGTCTATTGTTGGAGG	GCACAAAGATCTGAAAGATGTG	SORT1_ex17	GAAATAGGCAGGTACCCAGG	TTCTTAGCAGAACACAAAATCCC
DCLRE1B_ex4.1	TCTTCCTTGATATTCTCCCTGG	GACTTACAATGGGCACCACC	SORT1_ex18	GAATGAGTGGCTGCCTTGAC	TGCCCTAGATCAGCCTTCTG
DCLRE1B_ex4.2	TTGCTATCCTTCCCACAAGC	TCTTTCTTGGCCTTCTTTGAG	SORT1_ex19-20	TGTAAGAAGTGCACCCCTGC	TTTCCTGAAGTTGGAGCCAC
DCLRE1B_ex4.3	AGGAGGCTAAAGAGGCCGAG	TGGCTTCTGGACCTCCATC			
DCLRE1B_ex4.4	CCCCACTGGGATTTTCAG	TCCCACCAGCCATCAGTTAC	SPAG17_ex1	GTAGCGGGGACGCTTAGG	GTCGCCTAGGAGGGTCTGG
			SPAG17_ex2	GACACGAAATGACCTATTCAGG	CAAAACAGCCCAAGAATATGATAG
DDX20_ex1	CGGGAAAACTGCCAATAAAG	GAGGGCTCTTCCTGAGCTG	SPAG17_ex3	TTCTTGCTAGGTTCCCCAAC	ATCTCCAATGAGGAGCAGGC
DDX20_ex2	TCTGATTCCATGTTAGCCCC	TCTGACCTGTGGTAAAGTCCAC	SPAG17_ex4	TCATCACTGACAGACTGTGCC	GGCTCCATTTCATTCACCAC
DDX20_ex3	TCATGTAATGAATATACCAGTCCCC	AATTATCACAGTTATCAAAGACTCCC	SPAG17_ex5	AAAAGCCATTGCTTGCTCAC	TGGTCTCATTTTATTGCCACTG
DDX20_ex4	AATTCAGCAGAACAACCTGTAAG	CAATAATCACCACTTGGAAAAGC	SPAG17_ex6	ACTGGCTGAGTGCAACCTG	TGGCAAATCAAAGTCTTGAAAAG
DDX20_ex5	TCTTTAGCTTTTCCAAGTGGTG	ACTAGGCAAGGGAAAGCTCC	SPAG17_ex7	TTGCAATGCATATGACTCACC	GGAGCAAATAGAACCACTGTGTC
DDX20_ex6-7	TGATATGCTGGAGCTTTCCC	CCTCCTGACTTTGGAGGC	SPAG17_ex8	GGTTGGGATTCAGTGGTTTG	TGCCAGTAGGCAGTATGACC
DDX20_ex8	GGACAGAATTAACTGTGGTTGG	GAAAGTATTTTCCTCAGCACTGG	SPAG17_ex9	CATACCAGCATTTGATGAGGG	CTGTAGGTGAAGGGAGGCTG
DDX20_ex9	CCTCTGCTCCTCAGTATTCAGTTC	TCACATTATAGATCCCTCCCC	SPAG17_ex10	TGAAATTTTGATGGAAATAATGG	GCCAGGAAAGATGAAACGAC
DDX20_ex10	TGCTTTCTTTAAGGGGAGGG	ACCCAGCACATCATAAGCAC	SPAG17_ex11	ATGCAGCAATGAGCAGTGAG	TGCTTTGCTTTCATTTTCTTC
DDX20_ex11.1	GAAAACAGTGGTCTAGAGTAGCTTG	TTTTCAACGGGAGTCTGAAC	SPAG17_ex12	ATCGCAGCTCACATTGTGTC	CAGGCAAATTTCCTCGGC
DDX20_ex11.2	AGCTTCCTGTGAAAAGCCAC	GTCTCCAGATTGGCTACTACTTG	SPAG17_ex13	CCCAAATAATGAGGGTGAAGG	AAATAAGGGTGTAATTTTAAGAGGC
DDX20_ex11.3	TCATCAGGCACTACACAGGC	TCTGCAATCTGAAGGAAGACC	SPAG17_ex14	CCATTGTGGCGTGACTCTC	TCAGTGTAAGGAACATCCATGAG
DDX20_ex11.4	ACCAAATGGAAGTGACACCC	TCAGAAATGCCACTATACACCAC	SPAG17_ex15	CCAAGCAGAAAGCAGGTACAC	GAGAACTTATGGAACAACAATACTCAG

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
			SPAG17_ex16	AGCACAGCATTAACCAGCAG	TTTTCAGTGTTTCTTACTCCAAGTC
DENND2C_ex4.1	TTCACATTTACTTGTAATGGAACC	GAAGTGCTTCTATTTTCTTCCACAG	SPAG17_ex17	GTTAGTTGACGGAGCCCAAG	CAGCAGACTCTCCAGACCTATC
DENND2C_ex4.2	CAACTGGGTATGTTCTCGGG	TCACCTAAGCTGGGAAATGG	SPAG17_ex18	AAAAGGGAGTAGAGAAATGCCC	CCCCTAGCTTAGGGTAAGGAAC
DENND2C_ex5	GCCTATGCAACAGACTGAGATTC	TACACCAAACGTTTCCATGC	SPAG17_ex19	AAAGTTTATTTGCACAGTCAATATCC	TTTGGAATGCAGCTACTCTATTTAC
DENND2C_ex6	CTTTCTTTTGCCAGCTGTGG	CTGGACCCAGTTAAATCCCC	SPAG17_ex20	CCAGAGATTTGGGCCTTC	CACTATCCACTCCCACTGCC
DENND2C_ex7	AGCAAGACCCCATCTCAAAC	TCTCATTCATGGAAATGCTG	SPAG17_ex21	CAAAATGAACCTGAATGGAACC	CCGTATTTTCCCAAAACAATTAG
DENND2C_ex8	TCATCAAACTGTGAAACTTTTGG	ACATTAAGTTATATGGTGATTCTCCTC	SPAG17_ex22	CAGAAATTAAGCTTCTGGCAAC	GTAGAGGGCCCCTGGTTATC
DENND2C_ex9	TGGGTTTACATTGTTCTCATCC	GGTGCCACTCCACCTAACC	SPAG17_ex23	GTCGAGGAGCTTAATGTGGC	CATTCATTTAATGGAGTGGAGTG
DENND2C_ex10	GACCTCTGGATGCTGAAAGG	TTTTCAAAGCCACCAGATCC	SPAG17_ex24	GAACTCCATTCCCAGCCTTC	AGGGACCCTTGGCTGTG
DENND2C_ex11	GCTTATCAGCAACTACTCAGGC	ACATCCTTGACTTGCCACAG	SPAG17_ex25	CATTACTCCTGAAATCTATCCAACTG	CAAAGTTGAGACTACAAGGGACG
DENND2C_ex12	TTGGACTCTAGTTTTCAATTGTTTG	ATCTAGAGATAGTGTTTCTGAAGTTCC	SPAG17_ex26	AGAGACTCCTGCGTCCTTGC	GGCATGAAAATACACAATTTCC
DENND2C_ex13	TGAGAGCCTTGGTTACTGCTC	GGCACGTGGTTCCTAGTTTAC	SPAG17_ex27	ATGATAAAATCCAGCCCTGC	CCCATAAGTTGTCTTTTACTTTGG
DENND2C_ex14	TTGGCAGTTCCTGTAATGGC	CTGCCTACTAGTGTAAGTCCTGC	SPAG17_ex28	TTTCCATTTGACACACAGCG	CTGATGTGCAAGAATGTATAACAAC
DENND2C_ex15	ACAGACCCCGTGTAGCAGAC	AACTGAGAGACTGATCCTACATTTC	SPAG17_ex29	TGTGTCAAAACATAGTGAATGATG	CTCTTTTGGGTGTTGTTTCAG
DENND2C_ex16	GAAATGAGCTGGATTAAGATTGAAC	TGTGTACTATCGAATTTTCCCATC	SPAG17_ex30	GATGGTGCCATCTTCGTATTC	TGTATGTTCAGCAGTCAGTGAGG
DENND2C_ex17	TGGCTAACTTTTAAATGGGCTG	GGAACTAAAAGTTTTGCCTGAG	SPAG17_ex31	TAGGCTGAGGTCTGTTTTGC	CAGAGAAGTGGGTTTTCTCAGC
DENND2C_ex18	GTCACTTGTGGAATTTTGGAG	AACATAGTGCTGGGCATTCAG	SPAG17_ex32	AAATGGTTTTCCAAGTCCCAC	AAAATCTACATGGAAGCTTTGTAGTG
DENND2C_ex19	ACCAGAGCCAGTGTTTCCAG	CAACTCAGATTTTAAGCCCTTG	SPAG17_ex33	CATCTCCAATTTTGTGAGTAAACC	TTCCGTTAAAATTCTCCTTGAAC
DENND2C_ex20	TGTTTTAGCCCTGGAAATAATACACTG	ACCGCGCCTGGTCTTCAG	SPAG17_ex34	GAGTTCAAGGAGAATTTTAACGG	GGAGGATGAGAATTGGCAAAG
DENND2C_ex21	CTCTTGGGCTGTTCTGAATG	CATGGGATCCTGACCCTTAG	SPAG17_ex35	TCAATGCGTTTTCATTTTCTC	TCTGGTCCATTGCCCTTG
			SPAG17_ex36	TTATCGATTTTCCTTATATTTGCAG	TCTGAAAGACTGACATCCTTGTG
DENND2D_ex1	AGCCCCTTTCTCAGGCTTAC	CCCGCCTAAGAGGGTCAC	SPAG17_ex37	CTTTGCTTTTGCACTTCTCAG	AGGGAGCAAATGCTTCATGG
DENND2D_ex2	AGGTCACCCGGATTCTCAC	CACCTTCCAAACACCCCAG	SPAG17_ex38	TTTTACAAAAGATCAGCTTGAGAG	CACAACATGGCCGATACAAG
DENND2D_ex3	AAGCCCACCAGCTCTGAAG	CTCACACATGTTCCTTCTCCC	SPAG17_ex39	TGTACTGATTGTGTGTTGCCC	TTCCATTTCCTCAATCACTGC
DENND2D_ex4	TCCACCCTCTAGTGCTGACC	CCGTGGAGCTGAGCAGAC	SPAG17_ex40	AATTCATCATATAGTGGTGAGGG	AGGACATTTTGGCCTATCCC
DENND2D_ex5	CAGTTGGGAGAAGCCCTTG	AGCTTAGTTGCACAGGCAGG	SPAG17_ex41	TTTGAAGTAGGAAATGTGAGTGATTC	GGTGCAAGGCCTGAAATTAG
DENND2D_ex6	GAGTCTCTTCCACCTGCCTG	CACAGATGCCTACCCTCTCC	SPAG17_ex42	GTTTGTAAACTTTATTTTCAGGGC	TGCAGGTCACTGTAGGCTATG
DENND2D_ex7	GTAGAGGCCTAGCACCTCGG	GTCTGTTCCCTTGAACCCAG	SPAG17_ex43	TGTAAAACACCTTCTTCAATGTTTG	TAACTCTCAACCTCGCCTCC
DENND2D_ex8	ATGAGGGGTGTGTAGCATGG	GCAGAGCTGGGATCTGAATC	SPAG17_ex44	TGATCATGCATTATACCCTTGTAAC	AAGAGCTTCATGGTGAGATGG
DENND2D_ex9	AGACGGTAAATCCTCTGGGG	TGGATGAAACTAGAGTGGTAGAGTTG	SPAG17_ex45	TTCCTCTGGTGACAGGGAC	CAAGTCACAGGGTAGGGAAAG
DENND2D_ex10	CTTTTCTGCCAACCACCTTC	GATTGAGAAGCATGCCCTTG	SPAG17_ex46	CATCCCAGCTTATTCATTGC	GGTTTCAAAGCCCACTTCAC
DENND2D_ex11	GGTGAGTTTGGGTAATTTCATCC	CCTCCTCTTTATTCTAGAGGTAACCC	SPAG17_ex47	TTAAAAGCAAGGTTTGAACTCAG	CTTCCATGTTTTCCTCACTCG
DENND2D_ex12	ATTTCTGATCCTTATTCCTGAGC	GGGCTGAAGTCCAGAAATG	SPAG17_ex48	TGCCAGAGAAGAAATTCTATGATTC	GAAGAGGGGAAGAAACCTGG
DRAM2_ex3	TGGGGTGGTAAGTACAAGGC	TTCCCTTTATTGTTGTAATGCC	ST7L_ex1	CCGGTTTCCTGTCCTTCTC	CGAGAATCTCTTGCCCTTTG
DRAM2_ex4	GAGGTTTGAATTAACTCCTGATATTC	CTCCAGTGTTGAAAGAAGAAAATAC	ST7L_ex2	GCTGAATTACCTCATTTTGATTTG	GAATTGCTGGAACCTGGAAG
DRAM2_ex5	TTCTGGATAGCAGTGACCTTTG	TGAATGCTTCAGGTTTCCC	ST7L_ex3	AATTTGAGCTCATGTATCTGCTG	TCCCAGAGCCAAACTCAATC
DRAM2_ex6	TGTGTTGCCAATGTGTTCTG	GCACCAATAGTGTTTACTTAAATCTTG	ST7L_ex4	TCATCTTTGTACCATCTGTGGC	TGCTGCCAAATAAACTGTGG
DRAM2_ex7	AAGAAAATGGGTTTAGACCTTAGC	TCTACCAGAAGACTTTCATTGTTACTG	ST7L_ex5	TTTAAAACATCCAAATGTGTAGCAG	TCCAAGCCAAATAACTAAATAACTC
DRAM2_ex8	CAAGGGTTGATTTTCAGAAGC	AAAAGGCGGGAGAGTTGC	ST7L_ex6	AAAGCTTTGCTTACTCCACTACC	AAAATCATTTGCCTTGGTTTC
DRAM2_ex9	ACATTCATAATCACAAGGTAGCC	CTGTGAACCTTTCCCCAATC	ST7L_ex7	GCAAAGGGACATAACCAAGAG	ATTTCTGTCTTGCTGGCTTG
			ST7L_ex8	AGCAGCCCACAATACAAAAG	TCTGCATTAAAATAACTGGCAAC
EPS8L3_ex2-3	ATGGCTGCCTCTGGTTCAC	GTGTCCATGTTCTGGCAGG	ST7L_ex9	TGAATCTGGCTCAGCTCTATTG	TGCTAAATACTGACTGGTAAGAACAG
EPS8L3_ex4	GCTTTTCTCTACCCAGCACG	TAGTGCTTGCTTCCCCACAG	ST7L_ex10	ACCTCATGATCCACCCGC	AAAGCCCAAGAAGATCCAGG
EPS8L3_ex5-6	CCAGAAGCCACACCTAGCTC	CAAGATATGGGGTCAGGGTC	ST7L_ex11	TTTTGATTCAAGGGGCTG	ACCAAATTATGGTGACCCAG
EPS8L3_ex7-8	GAGCAGAATTGTGGGGAGG	GGTGTCCTTCCTTTTCCCAG	ST7L_ex12	TTCTGGTTTTCATAATTTGCG	CCAATCTGAACGACAAATCC
EPS8L3_ex9-10	CAGTGTGGGGAGTTCCAGG	AGGATTACTGGGGAGGAAGG	ST7L_ex13	TGTCTCCAAATATAGGATGGTAAATG	AAATGGAGAATACTATGACAAACAATC
EPS8L3_ex11	GTGAAGATGGGGTGGGG	CCTCCACCAGCTCCTTTCC	ST7L_ex14	TTGCTTAGCCAGAGACTTTCC	TCTTCCATTGTTCAGCTGGG
EPS8L3_ex12	ACTCTGGGCCTTTCCCAC	GATGTCCTGCCTGATGCC	ST7L_ex15	GGCCTAGGCTGTTCCTTTG	CCCTGTGAGGTAGGGGTTAC
EPS8L3_ex13	GCCTGGCATCAGAGTTCAC	GTGTGACCATGGTTTGGATG			
EPS8L3_ex14	GATACCCCAGCTCCCTCC	AGATCCAAAGGAACAGCCC	STXBP3_ex1	GCCAAAGTAGGTTGGGAGTG	GGTTGGGGCTGCAAAAC
EPS8L3_ex15	ATCTCAGGGCTGGGACATAG	CCTTGGCTGCCTGTAGCTC	STXBP3_ex2	TAGCCATTTTGGACCCTTTG	CACACATCCTCCCACATACG
EPS8L3_ex16	GAGGGAGAAAGCCTTGTGTG	CACCCCGACCTCCAACTC	STXBP3_ex3	TTTTGCCTACTGTTCTGATATTTTG	AAAATTAATACCTCAGATCTCCCC
EPS8L3_ex17	CTGTGGTTGGCTGTGAGATG	TGTATGTTCTGGGTCTCTGGG	STXBP3_ex4	TGAAGACAAAGAAAATGTTAAACCC	CATTCGGAAGAATGCAAATAATC
EPS8L3_ex18	AAAGAGCCCTGTCTCTGCC	ACCCAGGAGCACCAGACAC	STXBP3_ex5	CTGGCTGGAGAACATAGCATC	AAAAGTCGACAGTTTAAAACCAATG
EPS8L3_ex19	TGAGGCTCCAAGCTTATTGG	GATCTGCCATCTTGCATCAG	STXBP3_ex6	CCTCATAATAACTGTGAAAGCCAG	TGGATTAGGTTCAAAATTCACC
			STXBP3_ex7	TTAAATCAGGTCTCCAGATTAACAG	AACCAACATGCCAAGGTACAG
FAM102B_ex1	CGAGCTGAGCACCAAAAGG		STXBP3_ex8	AAAATCTGAAGCAGATATGATAAGG	GAATGAGCAAGTTTCTTTGGC
FAM102B_ex2	AGACTTGAGTGAAAGTATGGGTG	AAACAAAGCCAGCCAGTGAC	STXBP3_ex9	GGGAAAAGGGACAGAACTTAGG	GGCTTTCAGATTTCTGGTGC
FAM102B_ex3	AAATAGTTGGATTTAAGATCGTCAG	AAAAGAGCCAAGCTTGAACAG	STXBP3_ex10-11	TGTTTCTTTTGATTTTGGAGGG	GCAATGGGTACGTAGGAGTTC
FAM102B_ex4	TTGGAGAAAACCTGTCTAGGG	AGCACTGATTCACATCTTGTCC	STXBP3_ex12	TTATTGTGAGTGAAACTGGAAGC	CGAAACCCACTGCTAAATCG
FAM102B_ex5	GCTTTCCCAATAGAAAATATCGG	GGAAGAAACTATCCCTTAACTGG	STXBP3_ex13	TTCTTGAGCTGTTCCCTTGC	ACTGCCTAGGTGGCTCCC
FAM102B_ex6	GCTCACCAATCAATAACTATTGTAAAC	TTTAGGGCATCAAGTCGTCTC	STXBP3_ex14	CGTGAGCCACTGTACCCAG	CCAAGTTCAGCAACTATCTTTCAG
FAM102B_ex7	TTTGATAATGAACTTGGCAGAG	GCAGTTTCAAGTCCCTTTGG	STXBP3_ex15	TGAAATCGAGAGAAGACTGAAGG	GCTAGAAAACTGCATTTGCCC
FAM102B_ex8	CCAAAGGGACTTGAAACTGC	TGTTTCACTGGGCATCTTTG	STXBP3_ex16	GCACAAAATACTGTTCAGTGGAG	CCTCCAACCACTGCAGAAAC
FAM102B_ex9	CGCTGAGCCAAATCTTGATAC	GATGTTCTGAGCCTCTTTACTGC	STXBP3_ex17	AAGAAAGGGAGTCATGCAGC	TTGCTTATCTGGATTCATTTAAAAC
FAM102B_ex10	AACCTTTAGAGTATCTTTAAATTGCC	TGTGTTTCCTTGCTTCAGCC	STXBP3_ex18	TGGGCTACAAAGTGAGACTCC	GGAGCAAGCTCCTTTTCCTC
FAM102B_ex11	CCCATCTCTCCAGGGGTAAG	AAGTGCTTGAATAAAATGCCTTG	STXBP3_ex19	ATTACCCCAGGCTACTGCAC	TGTATTTGAAAAGCCATTATTCC
FAM19A3_ex2	TGCACAGGGGAGCTTCTG	CTGCTATGGAACACGGTGG	SYCP1_ex1	CAATTGTTCCTAGCGATTAGGG	GGTGCAGTAGCACAACATCG
FAM19A3_ex3	CTGCAGCAGGGAGGAGG	GGTATGAAATTGTATCCCCAGC	SYCP1_ex2	AAAACTTTCCTTCCCCTC	GAAAGTACACTTTTCTTCATTTCCC
FAM19A3_ex4	GCCTGGAGTAGGGGAAAGAC	CCCTTTTCTGCCCTTGGTAG	SYCP1_ex3	CACTGTATGTACATTAAAGGATTTGC	TCGATGTATGTAACACAATGTTTCT
FAM19A3_ex5	GGCATGCTGGAATGGAATAG	GCATGTAAGTGAATCCCCATC	SYCP1_ex4	GCTTTGAGGCTGCAGTGAG	GTCAGAATTACCAACCTAAAAAGATAG
FAM19A3_ex3	TGGCAGTCTCTTCCATAGCC	GCTCATGAAAGCCCTGAGAC	SYCP1_ex5	AATGAACTGCTTATTTGAAGACG	TTAATCAAATAATAATCCTTTTGAAGC
			SYCP1_ex6	TCAGTAGGCAGCTGTAGTTTGG	AACAAGATCAATCCAATTTATCCC
FAM40A_ex1	TCAGCCAAGATGGCTTCTAAC	CAACCTGGCCCCTAGAGTG	SYCP1_ex7	AATGATTATATTCGTGTTTATTTTTGC	TTCCTTTTTCCATAGCACTCAG
FAM40A_ex2	GAGGCAGTGGTGTCCTGTAAC	CTAGGGAAAGCAATGAAGCC	SYCP1_ex8	TGAATAATGTGATGTATGCCAAAG	TCCCTCTCCACCTCTACCAG
FAM40A_ex3	TCTGCACTTGTGAGACCTGG	AAACCTGCAGCTCTCGTTATG	SYCP1_ex9	TGTTTTCTGGAGGCATTTTG	AGCAATCGCTCTAATTTGCTC
FAM40A_ex4-5	GGGTCCATGCCATTGTGAG	TCTGCAAGCTCTGTCCCC	SYCP1_ex10-11	CAATCCATTTGCTGCTATTTTG	TTAATTCCCTGCCACATTCC
FAM40A_ex6	TGTCAGGGACATCACTTATCTTC	ACAGACAGCCTCCCACACC	SYCP1_ex11	AAGTGGACCAGCTGAAAAGC	GTCAATGTTTAGAACATAACAAGAGG
FAM40A_ex7-8	ACTAGGGTGCTCCTTTGACC	CAGTCTCCCTACTGCCCATC	SYCP1_ex12	AAGTGGACCAGCTGAAAAGC	GTCAATGTTTAGAACATAACAAGAGG
FAM40A_ex9	TTCTGAGCCTGACACAATGG	AAGCATGCACCTTGAGGG	SYCP1_ex13	AAAATATGAGGGTGGGAAAGG	TTGGCAATGCAAAGAAAATG
FAM40A_ex10	TCCTCTGGCTGCACTTCC	CAGGCCTGAGGGTGTGAG	SYCP1_ex14	GGGGCAGTGAATGTATAGAAGC	TGGCAATGCAAAGAAAATGG
FAM40A_ex11-12	GGCTTCACAGTTTCTGGATG	AAAGGCTTCTCCCTACAGGC	SYCP1_ex15	AAGATATATCAGTTGCTAAATTGCAC	CTGGCAAATGACAGAGGGG
FAM40A_ex13	TAGAGTGCAGGGAGTTTGGC	GGGGCAACCTTGTGGTC	SYCP1_ex16	TGTCTCTTTTCCTATTCTTTTAATCC	AAGTTAAGGGATTCAATATTCTTCTG
FAM40A_ex14-15	GGCTCAGGAGATTGTATTCTGG	GGAAAATGTTTGCCCTTCTC	SYCP1_ex17	CCATTTGGGCAGTTTGATTAC	TGCTTGGAAAGGAATTTTGTG

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
FAM40A_ex16	TCCGGAGTCCATCTTTTCTG	GAGCAGCAGTGCCCGAG	SYCP1_ex18	CGTTTAAAATATTTGACATAACCTGC	AGCAATGAAATTAGCACATTATTTAG
FAM40A_ex17	AGAGGGCAAGGTCCCAAG	GGACACAGTGGGAAATACCAG	SYCP1_ex19	CAGCCCTTCATCCCTTTCC	GGTCAAAGGGTCTTTAGAGGTC
FAM40A_ex18	GGAGCTGGCTCAGGCAC	TCTCTACCCACTGACCCTGG	SYCP1_ex20	TTGTTGGGAGAATAATTAAAGTAAGG	GGTTTATGATGCTTTAAAATTTCTGTC
FAM40A_ex19	GGACTTTTGGAAACTGCTGC	ACTCCCTGCTCCCAGCAC	SYCP1_ex21	TTTTTACATTTGGTTAAAGAATGTGG	ATTTTAGTTTAGACATATGGTGTTGC
FAM40A_ex20	AGTGTGACCTGGGATGTATGC	GGCAGGGAGTGAAAGATGTC	SYCP1_ex22	GAAGTGACCTTTATCAGTACTTTATGC	GAAAGCATTATTCTCTAAAACTGTGG
FAM40A_ex21	CATGGGGAGCTGCTACAGG	CACTGCAGCAATGAACTAGGAC	SYCP1_ex23	AGAGAATAATGCTTTCCCTTCC	TCAACTTTTCAATTGATACTGTTATGG
			SYCP1_ex24	CACAAAACCACAAAATGAATATGG	TTTTCGGGGTGTTTTATGAAG
FAM46C_ex2.1	TTTGCCATGTAGAAGTGAATCC	CCGTGCAAACCTTCACTAGC	SYCP1_ex25	TCCTGGTGTTTGAGGTTTGC	TGAAGAAATGGGCATCAAGG
FAM46C_ex2.2	TCTGAACTTCCTGCCAGAGG	TCCTCTGCTGTTCAAGGATG	SYCP1_ex26	TCCAAGACTGTTGTGTGTTTCAC	GAAGATCTATAAGATCCCTTTCAGAG
FAM46C_ex2.3	TACAGCAACCTTCTTGTGCG	CCTACCTGAGAGCCCTAGCC	SYCP1_ex27	TCAGAGCACTAGTTTATATTTTCATTG	TTTCTCATCCCGTTATTAAAATTTATC
			SYCP1_ex28	GAGAGGTTCGTCATCTTTCTTTG	CCTTCTCCAAATCTCCAAATG
FNDC7_ex1	CCTGCCAACATTCAAGTCTC	TTTCCAGAGTATTCAGAAACATTGTC	SYCP1_ex29	GTTGTGAATTCTACTAGTTGCTGC	GATGCCAGATATATGGTGATAACAG
FNDC7_ex2	AAGATGCGAAGAAAATCTCCTG	TAGGTGATCCATTCTGCTGC	SYCP1_ex30	TTGCCTTAAAATTATTTGTGAAACC	TGCAAAAGCACAAGGATCTG
FNDC7_ex3	CCAGCTGCATTGCTCTTG	GGTCTCACTGAATGCCCATC	SYCP1_ex31	CTCTTTCCTCCTCCCTTTGC	TAGAGATGAAGGGGCCTGAG
FNDC7_ex4	CTGTGATGATCTCACTGGGG	GAAGTTCATGTCTTGAAACACTCC	SYCP1_ex32	TCCACCAGCTTCTCATCTTTG	TTGGCTCTGGCAAATAAGAAC
FNDC7_ex5	GAAACACAGGCAAAATACCATTC	TTTTAAATGAGGAGGAAAGGC	SYCP1_ex32-UTR1	GACCGTTGGGCTGTAATTG	AAAGAATGACCCTCAATAATGC
FNDC7_ex6	TTCCCTGCAGTAATCAAGGC	TTTGATTGCATCTTAATTCACAG	SYCP1_ex32-UTR2	AACTACATATTGTCTGGAAACCTGTC	CAATGCTCTCAAGACTTTACAATG
FNDC7_ex7	CCCAAGTTCGTTCTTGTTCG	TAACCCCAAAATGGAGCAAC			
			SYPL2_ex1	CAGCCAGACTGGACTCCG	GTCCAGGGGTGGGATAAGTG
GDAP2_ex2	TCCAGTGAATATTGAATTGCTG	TGTATTTCAACAATGCTAGGGC	SYPL2_ex2	CCTTCCCAGGGTTCCCC	GCTCCATTGCTTCAGGTCAG
GDAP2_ex3	TTTCTACTATTGCTTAGCTTGTGAGC	TGCCTCTTTAAGGTGGAGGAC	SYPL2_ex3	CGGGGAGTCACGGTTAGTTC	CAGGTCCTGCCAGTCAAGTG
GDAP2_ex4	TCTCTGTTTCATGCTTTCATTTG	AACTGGGGCCTTTAACAATG	SYPL2_ex4	GACTACTTCTGGGCCAGGG	TCCAGTTCTGACCTCAACCC
GDAP2_ex5	TGAGTGGTACAGTTGGGTTAGC	ATGCAATCACAGCTCACTGC	SYPL2_ex5	AAGAAGCAAGGGAATCATGG	GGACCTGCTAGCTGGGACAG
GDAP2_ex6	CTGAATGAAGTGAGAAGTGATCC	GATCATTAGGATTTGTCTGCTGG	SYPL2_ex6	TCCCTGTTCCCACAATGAC	AGAGAAAGGTCCTGGAAGCC
GDAP2_ex7	CTGAGGTCTTCTTTGTTTTGTTC	CTTCTTTGGTGATTGGTGCC			
GDAP2_ex8	TCATCCTTACTGCAGCATTCAC	ATGCAAGGTTCCTTTCTTGC	SYT6_ex2	TATTGTGTGTGGCGTGGC	CACAGGCTGGAATCCCATAG
GDAP2_ex9	AATGCTTCTTTAACTCCTTCCC	CAAATAGCATGCTTGGTAGGG	SYT6_ex3.1	CCTTCCCCTGAGCCCTG	CATTTGCGGTCAGGCAG
GDAP2_ex10	TTTTATAAAGTCAGAGCCCAAGAC	TGTAGGCCTTGATTCCAAATTC	SYT6_ex3.2-	GCGTATCCTGAAGGCTTTTG	CACCCAACCCAGAACCAC
GDAP2_ex11	CTAGGGCATAGAGTAGGTTAATTTTC	CAAAGCAATGTTTGCCACTG	SYT6_ex4	TACTGCACTGGACATGAGGC	CCGAACACCCTTCCAAATC
GDAP2_ex12	GCATTCACCTAAAATTCAATCTG	TCTTAATACTCTGGGCATGCAG	SYT6_ex5	GCTTCGGTCCCTTTGCTG	GTGCTGGAGTGGGAGTGAG
GDAP2_ex13	AAAATTCAGATTGCGCTTTG	GTTGTGCTCGTAAGTGCTGC	SYT6_ex6	CCACACTGGACCTCCTTTTC	AGCTGACTGCCATTCTCCAG
GDAP2_ex13*	ATTGCAAAATTCAGATTGCG	GAGAAAATGCCTCTCTTGGG	SYT6_ex7	TCATGTGTTGTCACCTTCTGG	CTTTGGGTCCACACCAGG
GDAP2_ex14	GGCCTGCTTAGTGCCAGAC	GCTATTCGTGGAGGAAGTGG			
			TAF13_ex1-2	GATCCATAGCTTCGCAGTCC	ACACAGGGTCTGGTTTCGTC
GNAI3_ex1	GCAGTTTCCGTGGTGTGAG	GTTCAGGCCTTCCAAGCG	TAF13_ex1*	GCGTCACCGGAAGTGAGG	TGAAGTCCCCGAACTCTGTC
GNAI3_ex2	TAGGACCCGTGGTTTTCATC	CATCCTCTGAATAGCCATCCTC	TAF13_ex2*	CTGCTTCCTTTGGCTTTCTG	CACACACAATTGGGGTTAGG
GNAI3_ex3	AAAAGCACCATTGTGAAACAG	TCTTGTTGCTTAAATTCATTTCC	TAF13_ex3	ACGTGGGAACAGTTGGTCAT	GGATAGCTTATTTCAAAGGGATG
GNAI3_ex4	CTGGCCTGTCAGAAAAGGTC	AAACATTTCCTTAAGTGGGGAC	TAF13_ex4	AAGAGAAAATCTTCTACCTCACTTGC	GCTGAAAACTTTGTGTTTCTCCA
GNAI3_ex5	TTTGCTATGCACATGGTTGG	AAATTTTACCCTGATTAAGAGATGG			
GNAI3_ex6	CATTTCAGTTTAGGGGAAGGTG	AAGTTATTTTCCATTTCCTGGC	TBX15_ex2	CGGACTTCCTGCTCCTTCAG	TGGAAGAATGTAATGGCTCC
GNAI3_ex7	TGAATGCCATTTAGTGCTGC	GCCACTACCACTGAATACTCTCC	TBX15_ex3	ACCCCTGCTACTTCTCCCAC	AATTGTCTTGGAATGCTGGTG
GNAI3_ex8	TTGGGTTATGTTCCCTCTCC	CAAGAGACATCACTGTAGCACTATAAC	TBX15_ex4	TCTCAACTTGTTTCCTGGGC	AAGTGGCATAGAGATTCCTTATTTTAC
			TBX15_ex5	TGCTTTTCCTATAGTGCTTCAAC	TTGTTGTATGTCAAATTCCCC
GNAT2_ex1	ACCAGCTTTTAGGGGAGCC	TGAGGTAGAACCCACCAACC	TBX15_ex6	CCAGGCCAAGTTCTATATCCAG	TACACAGTGGCGGAGCATAC
GNAT2_ex2	ATCTCCTGGAAGGCTGAGG	CTCCACTGGCTGTACTGACTTC	TBX15_ex7	CAGTCCACCTCTGGATCTCTG	AGCTCCACTCTGTCCCTG
GNAT2_ex3	CCAAATTCCTAATCCCTTAGGTC	TTCTTACCCATCTTTCCATATAGTTTG	TBX15_ex8.1	TGGCTTGCATGTGACTTAGG	GAGCTGGAGGCAGCATTG
GNAT2_ex4	CAGAAGCCTCCTGGAAAGC	TCCCAACCAGAGAGAAGACTG	TBX15_ex8.2	TACCAACTCCTCCCTCGTTG	TCTGACCACGGAGACTCTGG
GNAT2_ex5	CCTTTTACGTCTCTTAGCCTCG	CCCTTTTCAGGATTCCACAC			
GNAT2_ex6	ACTTCACGGGGTGACCATC	AGAGAAAGGGTCATGGGGAG	TMEM167B_ex1	CGGCTAGCCACTCCTGATAG	AGAACGGCCACGTCAGTC
GNAT2_ex7	GAGCACTGGTCTCTCCAAGG	TTCTCCCTTAAGTTCCTTTGTGTC	TMEM167B_ex2	CCCCCTTGTCAAGTGTCAGT	CCCAGCTCACTGTACTCCCTA
GNAT2_ex8	CAAAAATAGGCCCCCAACTG	TCCCACTTTGAAAAGAACGTATG	TMEM167B_ex3	CTGGCCGACAGAGTGAGACT	CAGATGCTGTGGTGTCATCC
GPR61_ex2.1	GTCGCTGGACTAGGATGGG	AAAGAGGGCGTGGTCAAAG	TRIM33_ex1.1	CTCTCTCGCTAGCTCTCGGC	GGTGTCCAGGAGCGAGG
GPR61_ex2.2	CCCTCCGAAAATTTGTCTTC	TCCACTGGAGTGAACAGCC	TRIM33_ex1.2	GAGGAGGAGGAAGGCGG	TTCACCTTAGAGACCCCTCG
GPR61_ex2.3	GGTCTTTGCTGTCCTTTACTTTC	GCTGGCTTGAAGAAGCAGAC	TRIM33_ex2	CATCATGCACAATTTCTTTATTCC	AGCAGAACCACCCTAAATGAC
GPR61_ex2.4	GCAGGTGGAGAGTGTGGTC	TCCTTGCAATCATAACTGGC	TRIM33_ex3	TGTAAGGAAAAGGATTTTGTTGAAG	ATAGGGGCAAAAGAAGGAGG
			TRIM33_ex4	TGTATCGAAGCACATCAAAGAG	ATCACAGGAGTAGCAAACAAATG
GPSM2_ex2	TGTAAAGGACTGGATTGGCAC	CAACTGAAACTTGAAATGGTCATC	TRIM33_ex5	TCAAGATTGACTCATTTAGCTGTTC	AAAACAGATTTAAACTGGACCCAC
GPSM2_ex3	CAGCTGCCCTGAACAAGAG	AAAGGTTAGTTTTGAAAAGCGATG	TRIM33_ex6	CTCGGTTAAGTGCCCAGTTG	TTTTCCTTTATTTCCATACCCC
GPSM2_ex4	TCTTGGAAAACTGTACTCTGGC	TGCAAAGCCCTTTGAATG	TRIM33_ex7	TAGGTATTGTGCCCATTTGC	GTGGTGGTGGTTACATTTGC
GPSM2_ex5	CCGCAGCATTAGATGGTAGG	ACAGCAGATCCACGGCTTAG	TRIM33_ex8	ATCTGAGGCTGACACTGTGG	GGGGACTAGAAACAGCACAC
GPSM2_ex6	TTATGGCTGTAAGCTTCCCTAC	TTTGCAATAAGGAGACGCTG	TRIM33_ex9	ACTGACTCCATTCCCTGCTG	AAACTCTGACTTTTGAACTCTGCTC
GPSM2_ex7	GCCAGATCATTTCCTCCATC	TGCCAAAAGTTCTGCTACAATC	TRIM33_ex10	GGATTGGGGCTAGCGATAAG	AACACATCAATGACCTGCTCC
GPSM2_ex8	GACAGCAGAAAGCAGCAGAG	GGGATTTTAAGGAAAACAAAGG	TRIM33_ex11	AATAAGTTTGAAAGGGCCCAC	TCACTTAGACTGGAAGGAGACTTTTAG
GPSM2_ex9	TTCTTTATCCCTTTAGTTCCTGC	AGAAATAGGGTTCCCTCCCC	TRIM33_ex12	AAGGCTCAGCTTGATATTGG	TCCCTTTCTCTCTTCAAGTAGACC
GPSM2_ex10	AAAAGAATACAATTAGCTGTCCTATTG	GTTGAACTTCCCATGTTATATATTCTC	TRIM33_ex13	TGAGTTCTGTGACCGGTAGG	GCTGAGTGTTAAGTATAAAGGGGTTC
GPSM2_ex11	TGGGTTTGTGTATCAGATTTTAGAAC	GGCAGCATACGGTACTCATTC	TRIM33_ex14	GGCTTTATACCACATTCCCC	AAAGAGATATCTGGCTTGTGCC
GPSM2_ex12	AGAAGAGAGCTGGGGATTTG	ACAAGAAAACCCAAATGTCG	TRIM33_ex15	CGTGCCTGAATAATTTTGGG	GGCCATCTGCCCTGAAC
GPSM2_ex13	CATCCTTTACCTTTTGCATTTG	TTCAATTCTCAGAAATAAGCCAC	TRIM33_ex16	AAATAAACCAGGCCCCAGTC	TTCAGAAATCATCAATATACCATCAG
GPSM2_ex14	CAAGGCCGAAAAGATCTAGG	AAAATAAGCCCTTCCTGAATGAG	TRIM33_ex17	TTTAGTTGAGTATCTAGTAAACAGGGG	CGCCAGAGCCCATACTC
GPSM2_ex15	CTTTCAGAAATCAGGCTGGG	CCTCAAACGCGCATGTC	TRIM33_ex18	ACTCAAAATTTGCAAGCAGC	TGTTACTTTTGACAGGACTTTAGCAG
			TRIM33_ex19-20	GTCGCCAATCCTATGCAGTC	AGGCAGGATATTCCAGCAAC
GSTM1	SEQUENCING IMPOSSIBLE	SEQUENCING IMPOSSIBLE			
			TRIM45_ex1.1	TGGGCTTAATCACCTGCTAC	CATACAGGACAAAGGATGCC
GSTM2_ex1	CTACCCTCTCTGGGCCTCTC	AGAGGAACCGTCCCTAGAGC	TRIM45_ex1.2	GGGGAGACTCTGACACAAGC	TCACCGAGGGATTTGTCTTG
GSTM2_ex2	GCTCCCCTTGTGCAGAGTC	GGAGGCCTGCAGAGTCAG	TRIM45_ex2.1	TGCCCAACTATTCCCAACAG	ATGTCTTCCAGCTGCTTCAG
GSTM2_ex3	CACCTGTCTCAGGGGTCTTG	TGAGAGGAATTGGGCAGAAG	TRIM45_ex2.2	GAGTGGAGGCAGTGGCAG	AGCAGAACAAGCCACCAATC
GSTM2_ex4	AGGCCCTGGTCTCCTCTTTG	TTCACTCTCAGCCTCACAGC	TRIM45_ex3	GCAAGATGGAGATGTAGTCACTG	TGTTTCCAGTTTTCTCACTGC
GSTM2_ex5	CACAAGATCACCCAGAGCAA	ACCGCGGCTTCATTCTTAG	TRIM45_ex4	CAACTGAAGCCAGCTTGTGG	TGTGAAATGTCTCCAGATCCTAAG
GSTM2_ex6-7	AGCTGTGGGGAAGATGACTG	GCATCACAAAGGAGGGAAAG	TRIM45_ex5	TGTGCCTCAGACTCTGTTTAGC	GTGGCTGTGTGGTAGGGAAG
GSTM2_ex8	ACTTGAGCCCACATGGAAAG	GTGCTGGGATGCAGAGTCC	TRIM45_ex6	ATGGTAACTGCTGGCCTCC	GTCTGGGTCCTCTGGAAATC
GSTM3_ex1	CTCCTCCAGTCCATCCTACAC	AGATTCGCGCCGTTTTTAC	TSHB_ex2	TTTTCTTGGTTTCTTTGCCC	AGGAATGGACTTCTTCAGGG
GSTM3_ex2	TGCCCCAGTCTTCCTATCC	CTCGGCAGCTCGAAAAGG	TSHB_ex3	TGGGGCTAAGCAATTCTTTC	AACGCCTGTGTAGGTATGGG
GSTM3_ex3	CTAGGGGCCGCTTAACCAC	TAAGACGGGTTTGGGCTATG			
GSTM3_ex4-5	AGGTAGGGAATCAAGAGAGAGG	TGGAACAGAGATAACCCTTGG	TSPAN2_ex1	GCTGCTCGATCTCCACCC	TCTGTCTTCCCGTTCTCTCC

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
GSTM3_ex6-7	GCCTGTGCCCTGATTAACTAC	TCCTCCACTATGGGATCCTG	TSPAN2_ex2	CATCCTGTGCCTAGAGCCTG	CTTTGCACAGCCCCATCTAC
GSTM3_ex8	CGTGAATGTTGCTGTTCTGG	AAGAGCGCTGACCCCTTAC	TSPAN2_ex3	AAGCCTAGTTGGTCGTGGAG	GAGGCACTGACTCCCTTCAG
GSTM3_ex8.1	ACAACAAGATGGCCCAGTG	ACAGGCAGAGCAGCCTATG	TSPAN2_ex4	TGAAAGCGATTCAGCCATTAC	TGGGGATTATTTGGTTTCTTAAC
GSTM3_ex8.2	GCCCCACTTGACACTATTCAC	TTTTACAATTGTCCTCCAAAGTAGC	TSPAN2_ex5	TTCAAAGGTGTGGAAGAGGAG	CCTGGTGATTGGTGAGAACC
GSTM3_ex8.3	AAAAATGGACAAAATCCATGAAC	CGTTGCCACATGGTACTCC	TSPAN2_ex6	CAGGGTGCCCAGACTTGTAG	CAATGGCTTCCTAAGCTAGCAG
GSTM3_ex8.4	TCAAGAGATCGAGACCATCC	TGGAAATGCCCTTTAACTTCC	TSPAN2_ex7	AGATGGAAAGGATAGGATCAGG	AAGATGCAAAGTTTCAAGGC
GSTM3_ex8.5	GCAGTGAGCTGTGTTCTTGC	CCCTTCTCCTGAGACCTTCC	TSPAN2_ex8	ACCCAGGAAGTCTGCTGAAC	ACAGCTCCTGTGACATTTGG
GSTM3_ex8.6	AAGAGTCTGGAGGGAAACACC	GCAACATCCTAGCTTAAGGGTAAC			
GSTM3_ex8.7	GGGATGGATAGGCACATAGAG	CCTGGAGTATGGCAAGGTTC	TTF2_ex1-2	AGGCGGAAGCAGAATTGG	CTGTGACGCCCGACAGC
			TTF2_ex3	TGTGGGAATTACAAACCCAG	CCTGGCTACAACTCTTTCAGG
GSTM4_ex1	AGTGAGGATCCAGCAACCTG	TGGTGACTTTCTCTGCACCA	TTF2_ex4	AAATAGTTGCAAGATGTTTTCTTGG	GGTATTTTCAGGTGATGCTGC
GSTM4_ex2	GGGGTGCAGTGCAGTGTAGA	GGACACCCAACCCAATTAAA	TTF2_ex5.1	CTCACTGCCTGAGTGCAATATAC	TCTCTGCTCCCTCTTCTTGC
GSTM4_ex3	TCCACCTGTCTCAGGGATCT	CCTCTAATCTGGGGAGAGGAG	TTF2_ex5.2	AAGAAGCAAAGAGAAAAGGGAG	TACTCCTTGTTGAGGGGTCC
GSTM4_ex4-5	ACAGTGAGTGCCTGGTCTCC	AGGGAACCCTCAGAGGATCA	TTF2_ex5.3	CAAGTGGTAAGAGTCAAGATGTCC	AAAACAACATCATCTTCCTCCTC
GSTM4_ex6-7	AGTTCCAGCTGTGGGGAAG	CTGGGAAAGCATCTGAAAGG	TTF2_ex5.4	CTCCTGGAGGACCAGCG	TCTGCCCATAACAAGCAATG
GSTM4_ex8	CACGTGGAAAGGCTGTGG	GAGGCAGGTGCTGGGATG	TTF2_ex6	AGACCATAATTTCAGAGTTTGGG	TCAACACATTCTTTATTAAAAGGGC
GSTM4_iso2_ex8	CCTGTGGCTTAGTCAACTCC	GAAATCTTTGCTGTCTTCTGATTC	TTF2_ex7	TGGTCCCTTCACTTAGTTTGC	TGACTACTGTACCGGGTGGC
			TTF2_ex8	ACAGGGCCATTTGTAGCTAAG	GGCTGCTGGAGCACTGTC
GSTM5_ex1	CTCCTGGGCCTCTCAAAGTC	GTACCCGTCGCCTCCTCTAC	TTF2_ex9	CGGCATAGCCTCTCATTAGG	CAACACTACGTGTTTAGCACAATG
GSTM5_ex2	TGACGCTGTGTGTGTGTTTG	GAGCTCAGGGAAGGGACAG	TTF2_ex10	GCCAGTGGTTTCTCCTCTTG	CTCTTGGGTGGCTCAAATG
GSTM5_ex3	ACCTGTCTCAGGGATCTTGC	ACCTCTAAGGTGGGGAGAGG	TTF2_ex11	GAAAAGTTAAAAGCCATTTGGTG	CTCAGGGATGGGAGCAAAG
GSTM5_ex4-5	GGTCTCCTCTCTGCCCTTG	TCTGCTACAGACCTGACTTTGG	TTF2_ex12	TCTGCCTCTCTCATTGTCCC	CCCATTAAACCATCTGTGGC
GSTM5_ex6-7	CGGTTTTAGTTGTGGGGAAG	TAGCTCCAGGACTGGGAAAG	TTF2_ex13	TGCTATTTCCCCTTGTGTCC	CAAGATAATCTGGCACATACTAAGG
GSTM5_ex8	TGCAGCAAAGCTACTTGAGC	AGGAAGAAGGCAGGTCCAG	TTF2_ex14	TTACTTTGAACCTTTATTCCATTCTC	GACACAGTCATGTTCCTTCCC
			TTF2_ex15	CAGTGTCAGTATGGGGAATTTG	TCACTCCCAAGAAAAGACAGG
HAO2_ex3	GCATCTGAAGTGTTCTCACTGG	TGTCCAAGTTGCTCTCCACC	TTF2_ex16	CTGCAGGCACAGTTAAGCAG	TGAATCCAAATCCCCATTTTC
HAO2_ex4	CTCCCTGCTGCAGAATCATC	AGAATCTTGGAGCCTCTGCC	TTF2_ex17	GTAGTTCCTGGCTCTCTCCG	GTGAAATTGACTTCTCAGGTGG
HAO2_ex5	AAGATGGCCAGAGGCTACAC	TGTTCGATAAATGCCTGTGG	TTF2_ex18	CAAGTAACAAGGTTTGCTCCC	AGGAAAATGTGTGAGGAATCG
HAO2_ex6	TTTGGAATGCATCAAGTAGCC	TTGATGTGGACACACTCACAC	TTF2_ex19	ATGGCAGGTGAGGGGAG	TTGGCAATATCAAATTTTCTCAC
HAO2_ex7	TGGTCATATGCCAGGAAGAC	GCACAGAAATGAATGAGAAGCTC	TTF2_ex20	GCATTAAAGCCCTGCAGATG	CAACAAGAGCAAAACTCCGTC
HAO2_ex8	AGTGGATGGTCAGAATGTTCC	AGAGGGAAGGACTGCTGCTC	TTF2_ex21	CGAGATGTGTTACGAGACCAAG	GAAGTCAGTTCTTGCTGCCC
HAO2_ex9	CCCATGATGAGGTGAGTTGC	TGTGCTGTGAGTTTGTGATCC	TTF2_ex22	GGCCCATAGTTTGCCATTTC	TGAAGGTATGCCTCAGGGG
			TTF2_ex23	ATATTATTAGATGTGCTGCTTAGGG	TTAGGTTGTTATTCCCAGATCC
HBXIP_ex1	CACGAACTGAAGCACTGAGG	CTTTCGTTCCGCTCAAGTTC			
HBXIP_ex2	TGCCTGGTACATAGAACACCTG	ACGCTAGGCTCCAAGTCAAC	UBL4B_ex1.1	CCCTAATCTTCAGGGCAGC	GCATCCTGTGGTTCAAAGTG
HBXIP_ex3	AAACAGCCAGCTAGCTTATTTTAG	AATCACTGATGAACACAGGGC	UBL4B_ex1.2	ATCAATGTCATCATGCAGCC	GGAACACGTAAGGATGAGGC
HBXIP_ex4	TTGTATCCTAAAGCAAACAACATC	ATGGCCTAACTACTGGAACTTTAG			
			VANGL1_ex2	CCAGACTCCTTGAGGTTTTAGG	TAGAGAACCCCACCAACCTG
HIPK1_ex2.1	TTGTGTGTTTTAAATCTAAGCCTTG	GATCTGCACACTACCGTTGC	VANGL1_ex3	GAATTTCCCTAAATTCAAAATGATAC	AAACATTTACAGACAGGAGTCTTGG
HIPK1_ex2.2	AAAGCAGCCAGACCCTGAC	GAAAGCACTCGTATGAACGG	VANGL1_ex4.1	TGTTCACTGCCCTTTGTAGC	GCACGAAACACAAACACCC
HIPK1_ex2.3	CCCCTCCTATGCCAGACAAG	GCTATATGTAGCCTACGGGGTATATG	VANGL1_ex4.2	CCATGGCATTCAAACTCCTC	CTGTCTCCCAGCAAGACATC
HIPK1_ex3	CTGAGCTTCAGATGATGGATTG	GCAATTGGACTACTCTTTCCC	VANGL1_ex5	TTGTGTTCCTTTACCTGACTCC	ATAGTCACCCAACTGTGCCC
HIPK1_ex4	TGTAAAAGCCTTTAATAATTTCCC	TTCAAAATTGGGGAGTGAGG	VANGL1_ex6	GGGGTGGGTAGACAACACTG	TGGAACCTGGGAGGCAG
HIPK1_ex5	CATCTATGGTTGGCAAGGTG	AAATGTTCCTTTACTACATGGTGG	VANGL1_ex7	TCTGTTCTCACACCCTGTGG	GCCAAGCTCAGTGACAAGC
HIPK1_ex6	GGGAAAGCAGCTATGTAGTGTC	GCCATTGATAAGGCGGTAAC	VANGL1_ex8	CAGTACTGCCCTTAAAACAGGC	AAAAGACACCCCTCTGGC
HIPK1_ex7	CCATGTTCATCTTGAGTTATTTGC	CCCTGTATTCTTACTAGTTTCTATCCC			
HIPK1_ex8	AAGCAGAGAGTTATTTCAGAGGC	AGCTGCCTAAAGCAGGTAGG	VAV3-iso1_ex1	CAGTGGTGTTTTACCACTCTGC	GCCTGGTTAAGTTAGCAAACAAC
HIPK1_ex9	AATCTTAAAACAGTACTCCCAGACC	ACCATTTTCTAGGACAAAATATCTG	VAV3-iso1_ex2	TTGCTTTCCAAAGTGCCAG	AGGCCAACAATATGCCTCAC
HIPK1_ex10	ATCCAGATATCCTGCCTCCG	TGCCCTCTGCAAAGACTAAG	VAV3-iso1_ex3	TTAAAATGTTTACTTGGACTTGGTG	TTGCTTTGCACTGTGGACTC
HIPK1_ex11	TTGTAGCTTGAGCCCCTCTG	ATTACCAAGTGCTCCCACCC	VAV3-iso1_ex4	GCATCATGGGGTTATCATTTG	GACCTGGCCAGTATCCTTTG
HIPK1_ex12	TTTGATGTCTGTTTGGTCCTG	CCAAATGCAAGGCAACATATC	VAV3-iso1_ex5	TCAGCTGAACTGAATTAATGATTG	TCCTGGAGCTAAGGGTATGG
HIPK1_ex13	GGGGAAAAGATACACAATTTGG	GGGCCCACTTGCCTGAG	VAV3-iso1_ex6	AGCCTCTGTTCTCTTCCAGG	TGTCAGATGGTCTATGCAATTATG
HIPK1_ex14	CAAGACGCTGTCACACACAC	TGGCAAAATATTTAAGCATTCAC	VAV3-iso1_ex7	GCTGTCACCAGTGTATTGGC	CCTGCAAACAATTTACAATGAAG
HIPK1_ex15	GGACTAATTAATATGATGCCATCTG	GGTTCTCCAGAAAGTAACAACAG	VAV3-iso1_ex8	GCTCAGGAGTTCATGTGTTTG	CTCATGCATTAGCGCATCC
HIPK1_ex16.1	CTTAACTTGGCCTTTGGTGC	GGGTGAGTGGTATAGGCTGC	VAV3-iso1_ex9	TCATCTGATTAGGTTGTTATGGTTG	TGCAAACAACTGCTCAGTCC
HIPK1_ex16.2	TATACGTATGCTGCCCCGAC	CCCAATATTAAGAACGCAGGG	VAV3-iso1_ex10	CAATGGACATAAAACCCCAG	TTGAACAGCCAGAAATGCAG
HMGCS2_ex2.1	TGATTGCTATGTTACTGTCTTTTGAG	TTCCAGCCTGCCCACAG	VAV3-iso2_ex1	TCCTGCTCCCTCGATCC	CAAGCTCAGCGCACCTAGAC
HMGCS2_ex2.2	TGCTCAGTCCAAGAGGACATC	GGAACTGAAAAGCAAAACTGG	VAV3-iso2_ex2	TCTTGAGTGTTACAGAGGGTGG	AAACAATTTGCTTAAATGCTTTCC
HMGCS2_ex3	CATAAGCTCCACAGGACTGG	ACCATCCTCAAACGCATACC	VAV3-iso2_ex3	TGCAAATTTATGTTTCTGCTGC	CATAGTACCTGTGTGCAAGATGC
HMGCS2_ex4	TTTGTGCTCAAGGTAGGCTG	AGACCATCTCATGGCCTTTG	VAV3-iso2_ex4	CAGCCTGATTGTTGCTCG	TCCACAGTTAAAACCCACAATG
HMGCS2_ex5	TATCTAAATGGCTCGCCCC	AGCCTCCTTCTTGCTCACAC	VAV3-iso2_ex5	GAGGGACAAAGTTGTAATGTAAACC	TGTCTCCAAATGCAGTAACAAG
HMGCS2_ex6	CCTTGCTCCCTGCCTCC	TTTGTTGACCCTGCAGCC	VAV3-iso2_ex6	GCAGTTAGATGTCTTTCAATTATGTG	TGTAAAATGAATAAATCTGCTTTGAG
HMGCS2_ex7	ATCGGGATGATAGGCTGTTG	CCTCAGTAATGGTGGGGAAG	VAV3-iso2_ex7	GATCAACAATTGACTTTCAAAACG	AAGAGTATTAAAATAGGGATCTGGAAC
HMGCS2_ex8	CAGAACCTGGCTCTCTGTTG	TCCCAGCTACTGGAATGG	VAV3-iso2_ex8	AAAAGGATTCTGGGTTTGGG	TGCTCAGGTAGTAATGTAATAAACAGC
HMGCS2_ex9	ACTGTGAAGGGGCAGACATC	TTCTTCTCCCTGCACCTGTC	VAV3-iso2_ex9	CCCCTCCAGTTCATTCTTGG	TGGGAAATGTACTGCTTCACTG
			VAV3-iso2_ex10	GGGCCTTTGCATCTTTTAATG	GGCCCAACACATGGAATATC
HSD3B1_ex2	GGAGGAAAATGAGGCATCTG	TGCTAGACAAGGTCCATCTCC	VAV3-iso2_ex11	ATGGAAATGGAGCCCAGAC	TTTGAATTCCATTGTCTTTAGGC
HSD3B1_ex4.1	CATAGATCTGTGTTCGTGGTTG	CCCTCAAGGCCAGAATGTG	VAV3-iso2_ex12	TTCTTTCAAAGCAGTTTGCATC	AGAATGTCGTTGATGTGGGG
HSD3B1_ex4.2	CCTGTACACTTGTGCCTTACG	CAGATCTCGCTGAGCCTTC	VAV3-iso2_ex13	GGATTACAGAGGACAGACTTATTTG	TCATGGGTTGGTCTGTCTCC
HSD3B1_ex4.3	GATTCCAGATGGAGCTTTCC	TGTGCCCTTGTCACTTTCTG	VAV3-iso2_ex14	CCCCGTTGTTTTCAGTGTATG	ATCTGGCTGTGGCATTATCC
			VAV3-iso2_ex15	TTTGGATAGTTGTGGTATTCTTGTG	AAAACATTCCCGTCGCTG
HSD3B2_ex2	TTTAGCCCTCTTCTGGGTCA	TGTGACTGGCTTTCATTTGG	VAV3-iso2_ex16	CAGATCTCTACCTGATGTCATTTTG	TTGGCTAGAAGGTTAGAAAAGTCTG
HSD3B2_ex3	CCCTCCACTCTAATACCCACAC	TCAGCAGGTGTTCAATGGAG	VAV3-iso2_ex17	CAGTCTGGATCCCACAAACC	TTTAAAACAAAGTTATGCAGAAAACTC
			VAV3-iso2_ex18	CAGTGGTGTTTTACCACTCTGC	AAAGCCCAAGCCTGGTTAAG
IGSF2 (CD101 in hg19)			VAV3-iso2_ex19	TTGCTTTCCAAAGTGCCAG	AGGCCAACAATATGCCTCAC
IGSF2_ex1	TTGTCACTCAACCTCTGAATGTTAG	CTGTTCCATCACTCAGCACC	VAV3-iso2_ex20	AAAATGTTTACTTGGACTTGGTG	TTTCTTAATCAACTTGCTTTTGC
IGSF2_ex2	GGTACACTGGGAACTGAAACAC	AAGAAAATGGAACAATATAAAGCC	VAV3-iso2_ex21	GCATCATGGGGTTATCATTTG	GACCTGGCCAGTATCCTTTG
IGSF2_ex3.1	AGGCCCACTGGACAAGC	AGCCTGAATGTAGTGGGTCC	VAV3-iso2_ex22	TCAGCTGAACTGAATTAATGATTG	TCCTGGAGCTAAGGGTATGG
IGSF2_ex3.2	AGGAAGCCAAGCCACTGAG	AAAACATACCTCATAAATACCCACTAC	VAV3-iso2_ex23	AGCCTCTGTTCTCTTCCAGG	TGTCAGATGGTCTATGCAATTATG
IGSF2_ex4	TGAAAGGACAGAGGCTGTGG	AGGGGTCTTCCACTGACAAG	VAV3-iso2_ex24	GCTGTCACCAGTGTATTGGC	CCTGCAAACAATTTACAATGAAG
IGSF2_ex5	GACTTGTCACAAACTAGTCAAGTCC	AACACGCAATCCAGTTCCAC	VAV3-iso2_ex25	GCTCAGGAGTTCATGTGTTTG	CTCATGCATTAGCGCATCC
IGSF2_ex6.1	GGTCCTAGTCAAAGAGGTGGC	TGTGAACGAAATAAAGACTGCG	VAV3-iso2_ex26	TCATCTGATTAGGTTGTTATGGTTG	TGCAAACAACTGCTCAGTCC
IGSF2_ex6.2	CCAGCCAGCTAGCTCTCAC	GGGAGGATTGCTATTTGTTTG	VAV3-iso2_ex27	CAATGGACATAAAACCCCAG	TTGAACAGCCAGAAATGCAG
IGSF2_ex7.1	ATATGATGGCGGGAGGATG	TCGCTGTCTTCCACATTCAG			
IGSF2_ex7.2	TTCCACTAATGCCTCTTGGC	CAAGAGAAAACAGCCGTCAG	VTCN1_ex1	TTAAGGCCAATACACGGGAG	TTACAGGTGGGGTCCTATGG
IGSF2_ex8	AAAATGAATGCACTGATGTCTAAAG	GAGCAAAAGATAGGTGAGAGGG	VTCN1_ex2	GCAAAGCAGCACATCAAGAG	TCCATGCTAACGCACTGAG
IGSF2_ex9	TCCCAGGAGTGTTTGTAATTG	ACCAGGATTCATTCACACCC	VTCN1_ex3	AGCAATGAAGGGTTTGGTTG	CAGAATTGGGAGCATCATTG
			VTCN1_ex4	AGGCCTAGCGCTTCAAAGG	TGGAATTGAGTGATCAAATTCTTAG

Name	Forward primer	Reverse primer	Name	Forward primer	Reverse primer
IGSF3_ex2	AATGAGAGATTAGGAGCTGGG	CCACGAGAAATTACAGGAACC	VTCN1_ex5	TCATGTGAAGAAGACATCCAGC	CATTAGGAGCACAAGCACCC
IGSF3_ex3	TCTTGGCATGATCAAATTCC	CCAAAGGCCTATCCTTCTGA			
IGSF3_ex4	TGCATTTCCAAATCCAAACA	GAGGTGCCCAAAGAAGGAGT	WARS2_ex1	CTGGGATTGGCTCTCTGAAC	AAATAGAGGGCAGTGGGATG
IGSF3_ex5	TGACCTGGGGGTCTCATCTA	GTGTATGCACAGGGCTGCTA	WARS2_ex2	GGATTCATGGGATTCTGCC	AGCAGGGGCTGTATGAACC
IGSF3_ex6	ACAGCGGTCCTCTTGTCAGT	TGGGTGTGGAGTTAGAAGCA	WARS2_ex3	TCTTCTTATATGCAAGAAAACACTG	CAGGATTCCCAGAGCCTAAC
IGSF3_ex7	TTGCAGCAACTAGGGTCAGA	ATCACTCTCTCCCAGCCAGA	WARS2_ex4	GGCTTCAGATCACAATTTGC	TGCCCATGAACCAAGTCTG
IGSF3_ex8	GGGAAGGGTTAGGCTAGTGC	GTGTGATGATGGTGCAGTGG	WARS2_ex5	TTGGTGAACATTTGTTGAATACC	GTCCCAAGAAAGTTCCAAGC
IGSF3_ex9	TTCTCTCTGGCCCGTAGGTA	CAGAAGCAGAGTCAAGGATGC	WARS2_ex6	TGGATATCACATTTGCACCA	TATCTCCTCTGGGCTGTCTG
IGSF3_ex10.1	GCGTACGAATTGCTTGTTG	GGAAGGACTCTCCAAATGC			
IGSF3_ex10.2	TGAGCGGGAGACTGTGG	CAGTATACTTAAATGGGAAACACC	WDR3_ex2	GTAGAACCAAGGCACCCAAG	AGTGGCGAAGAACACAAAGG
IGSF3_ex11.1	TGGATGTTACTATCATGTGTTTAATG	GAGCCTCTGGAAGCGAAG	WDR3_ex3	GCTTTTCCCTTGAGTTCATGC	GCTTCAGCATAAGCCTTACATTC
IGSF3_ex11.2	AAAAGGAGCAGCCCTGG	GCTCAAATTTATCTTCCAGACAC	WDR3_ex4	CATTGTTTGTCAAGTTGGGC	TGAGTAAGGAATAAAGCAAATTCTCTC
IGSF3_ex12	AGGAAATGTGAGCAGTCCAG	AGAGAAAGGGAGAGCCTCAG	WDR3_ex5	CAGTGAGTTTTGGATCTCCC	CCACCACCTTTTCCCTTTTG
			WDR3_ex6	GGAATGTCATATTATGCTACCTGG	AAATCCAAGGTGCTTATGAAAC
KCNA10_ex1_1	AAGGGCCAGTTGTATTCTGC	TTTAGGACCACTGGCTCAGG	WDR3_ex7-8	AACAACATGCCTCCTGGAAC	GTGTTTGGAAGGAACAGCC
KCNA10_ex1_2	AGGCTATGCCACAGACTTCG	CTGGTCCATGGCCTCACTAC	WDR3_ex9	GCAAACCCAGAGAGATTAAGTTG	TCAGTAACATGGGGCCTACC
KCNA10_ex1_3	GGATGCAGTTCTTTGACTCCA	CACGATGCAGGTAGACTCCA	WDR3_ex10	ACAGTTTGTTGGTCAAGGGG	GATGTCTCATGACCACCGTC
KCNA10_ex1_4	ACCATCTTCTGCCTGGAGAC	TGCACTGGAGAAGAGGATGA	WDR3_ex11	GTGAGCCACTGCACCCAG	AGCTCCACAGGCTAATGGC
KCNA10_ex1_5	TCTTCCGCATCTTCAAGCTC	GAACAGCCACCATTGGTCTT	WDR3_ex12	CCTCTAAAATGGGGATGCAG	TCCAGTCCATTTCCAATATGC
KCNA10_ex1_6	CATTGCAGGGGTCCTCAC	GAGAGAGGAGGGGAGAAAGC	WDR3_ex13	ATGAATCATGGAGGTTGTACG	GAAATCAGGGTAGCCACAGG
			WDR3_ex14	AGGGGCTAGGGAGAACCAC	AAAGGCCATTCCATAGCTCC
KCNA2_ex2.1	CAGTAGCAGCACCCTGAGC	GACGCTCTTCCTCCTTGATG	WDR3_ex15	TGTGCTCTAAAGAGTTAGGATTTGTC	CCCAAATTCTTAAGGTCAGTCAG
KCNA2_ex2.2	AACCGCCCTAGCTTTGATG	GACAGGCAAAGAACCTCACC	WDR3_ex16	CTGCCAAGCTGTCTTTTGTG	GCCAGAATTAAGAAAGTATCTGCTG
KCNA2_ex2.3	TCCGGGATGAGAATGAAGAC	CATCGGCCTCTGCAAAATAC	WDR3_ex17	TTTAATGTGTTTGAAAGAAATGGG	TGTATCACTGCCTGAGAACACTG
KCNA2_ex2.4	TTGTCCAGACACTCCAAAGG	CCTGGATCTCCATGTAATCAGAC	WDR3_ex18	AGCTAGTCAAAAGCTATTTATGTGTG	AAAGAATACCTCATTCATTGACCAC
KCNA2_ex2.5	CTTACCGGTCCCTGTCATTG	ATAAGCCGGTGATGAGGATG	WDR3_ex19	GGAGGTAACCACATGCCTTC	GGACTTGGCAATGCATCAG
			WDR3_ex20	AGCAGCTTCTCGCTTCCTC	GCTGGTGATACCTGCAACTTC
KCNA3_ex1.1	CCAGACCCAGACAGAGCATC	TTCAGCTGCGTCTCGAAG	WDR3_ex21	AAACATTTGCAGGATGTTCTTAC	CAGGACTACTCCCATCTTTGG
KCNA3_ex1.2	GCTGGTGAACCACGGCTAC	GTCCTCGCGGAACTTCTCC	WDR3_ex22	TTCAGTGTCCTGGGATGATG	TCTATTCCATTTATTGCATATCCTG
KCNA3_ex1.3	GTACTTCTTCGACCGCAACC	CAGCGTCTCCACCACG	WDR3_ex23	AACAGGTTTGATAATTCTTTTACACTC	CCAATAATCCATCCCCAAATC
KCNA3_ex1.4	TTGTCATCTTCTGCCTGGAG	AAGATGAGCAATCCCAGCTC	WDR3_ex24	AAGTAAGAAGGAGGGGTTCCTG	AGCACTTCATAAATGTTTGCTG
KCNA3_ex1.5	GACAGGGCAATGGACAGC	TGCATGTACTGGGATTGCTC	WDR3_ex25	CAGATGAATATTTAACTCTAGGCAAG	CCACCGAGTTCGGCATAC
KCNA3_ex1.6	CAGTTCCCGTGATTGTTTCC	AAACAAGGGCATAGGCAGAC	WDR3_ex26	GAAGGGTTAAGTAGGGCTAGAACC	CCCTCATTAAGAACTTTTCTTGTTTG
			WDR3_ex27	CGCTGCTATAATGAATTGAAGTG	GACGCCAACTCTTCTGTGC
KCNC4_iso1_ex1.1	GGGATAGGCAGGGGCAAG	AGTTGAGCACGTAGGCGAAG			
KCNC4_iso1_ex1.2	CACGCGACATGAGACCTACC	GGGCTCTCGAAGATGTCG	WDR47_ex2	CCAGTAAACCTGTTATGTGTAGTTGC	ATCTCGCCACTGCACTCC
KCNC4_ex1	GAAGGAGGAGATGGCCAAG	CTGGGACTGGGTCCAAGTAG	WDR47_ex3	TGCTGTTTTGTCATCTGTTCTG	TTCAAGGCCAGTGGAGAAAG
KCNC4_iso1_ex2.1	CTTCCTGGTAACCTTCATGC	TTGAAGATACGCAGGATGC	WDR47_ex4	TTCAGTTATGGCTTCAATGTGAT	ACACCAGATCTTGCCCCATC
KCNC4_iso1_ex2.2	ACCTACATCGAGGGCGTATG	GGGGTACATGTCTCCGTAGC	WDR47_ex5.1	TTCCTATCTTGAAAGCTGTTGC	CAGCATATGCAGCTTTTGTAGG
KCNC4_iso1_ex2.3	ACTACGCTGAGCGCATTG	CAAGGTCTGAGGTCCATTTG	WDR47_ex5.2	GCCTGCTTTATGAATGCTGTG	CCAGGATGGTCTCAATCTCC
KCNC4_iso1_ex3	AGCAAGGCTCAGGTCGC	GACCTCAGGCTCTTGGTGG	WDR47_ex6	TCGGGGCTACAGAGAGAGAC	TCCCAAATATCAGAAAATTCTTTTTAC
KCNC4_iso1_ex4	CTCTGAGGGGTTGGGTTTC	CCCTGGGATATTCCATTGTC	WDR47_ex7	CACATACTGTTTTTCAAAGAGCAA	CCCTCGGAATGCTGTGACTA
KCNC4_iso2_ex4	AATGTTGAGCCGAAAGCG	AGTTTCCCTAAGCCTCTGGC	WDR47_ex8	TTGGTTATGTCTTAAGTTGATGGTG	GCTCCAAAGACAAGTTCAACC
KCNC4_iso3_ex2	CCAAGAAACGGAAGAAGCAC	AATATACAGCCCCATGTACTGC	WDR47_ex9	GGAGTCAGCTAGGATTGTTTGC	CACCACGTCTGACTAAAATGG
			WDR47_ex10	TGGCAGTAAACCTAACATGATCC	GAATGGCAATTCTGAAAAAGG
KCND3_ex2.1	ACTTTGGCCGTTTCACTCAC	CGTGCGGTAGAAGTTGAGC	WDR47_ex11	GACACGCTTTTTCTCCTTCC	TCAAACACCACCTAGACACCTG
KCND3_ex2.2	CTTCAACGAGGACACCAAGG	CGTGTCCAGGCAGAAGAAG	WDR47_ex12	CATTATTCTTGACCAGTTTTGAGG	GTGATTGTGCCACTGCACTC
KCND3_ex2.3	GCTGGCCCTGGTCTTCTAC	AGCCCTTCTCGGCATAAAAC	WDR47_ex13	AGCCACCTTGCCCAACTC	CAAATGCAAGTATTGTGAAAGAAC
KCND3_ex2.4	GCCCTACTACATCGGTCTGG	CCTCCTCTACCCATGGTGAC	WDR47_ex14	GGGCCTTAAAAGGACACAAC	AAACCTAACAATAAGACTATTCACTGC
KCND3_ex3	CTGCCTTGTTTCCCAAATG	AGTCCTGGCTCCCTGACTG	WDR47_ex15	CGAGAGAGAGAGAGAGGCAAC	ATAAGGGGCCTCTTCGTG
KCND3_ex4	AAGCCAGCCTCACAGCTTC	AAAACAAGCCCATCTACCCC			
KCND3_ex5	TGGAGAAATGAAGGAGCTGG	GCCTTAGAAAAGGGTCAGGG	WDR77_ex1	CGGAGGCTCAGAGGTGC	GTCAGGATAACATGCAGGGC
KCND3_ex6	AAAACTGGCTTGGGAATGC	GGGGAGAATCCACAGACTCAG	WDR77_ex2	CATGGAGGCGAACCCTTC	ACCCTACGGTACCCCAAGTC
KCND3_ex7	AGGGGCTTACCTCTCTCCTC	TGCTGCAATTGTCTAATTCTGTG	WDR77_ex3-4	CTAGAAGTGAATTACTGCCTTGG	TCAGAACATGGAATTGCTGC
KCND3_ex8	GGTCAAGGAAGCTTTGGGAC	CAACATGCCAGTCCCCTTC	WDR77_ex5-6	TGATTATTCCTGGGCTGGC	CACTGGGCATCTAGCAACTATG
			WDR77_ex7	CAAATAAAGGTGCTCCCAGG	GCCTTGTGTGTTGTGGAGTC
KIAA1324_ex2B	GCAGGTGAGCTGCAGAGG	GTCTGAAACTCTCTGCGCC	WDR77_ex8	TTGCAAAAGAGAAAACTGATGC	AACCTCAGCTAGAACACTGTCTG
KIAA1324_ex3	GGTGGGAAGCTGTTAATGTTTG	ATTGGATGGCAGGAGACTTG	WDR77_ex9	GGTTTTGTCTTGGGATCCAG	AAGAGGAAAATGACTCGCCC
KIAA1324_ex4	CCATTCACTTCTGGGGCTG	CTGTCCAGTGTGCCCTGTAG	WDR77_ex10	ATGTGCCCCAGATGTGTTG	TCAGGCTGAGAGACGATGC
KIAA1324_ex5	GCAGAACATGAGACCTGCAC	GGATCTGAACCCCAGAGAGG			
KIAA1324_ex6	GCAATGTGGGGAAAGGTATC	CTCTCTGAGGAAAGGCCCTG	WNT2B_ex1	CCCTTCCAAACCCTGAAGAG	CAAGCTTGACCATTATACATACTACG
KIAA1324_ex7	CCCAACAAGAGGCTATACCAAC	CAAGGATTCAGCCTGGAGAC	WNT2B_ex2	CCTCTTCCTGCTGAAAATCC	TCTGAAATGCTATCTGATCTGTTG
KIAA1324_ex8	GCTGTTGACTTTGCAGTCTCC	CTCCTTGTGTGTTTTGGTGG	WNT2B_ex3	CGTACCCCTTCCTTATTCCC	TTCCCACTCACCACCCC
KIAA1324_ex9	GAAGAGGCCAGGAGCTCAG	AACCTCAGGGAGTGGGAGAC	WNT2B_ex4	TTGTATGGGCTGAAGAAGGG	GTCACACACACTGGTCACCC
KIAA1324_ex10	TCAAGGTAGAGTCCCAAAGATG	CTGTTAGTGCCCACACAGCC	WNT2B_ex5	TCCTTTTGGACCTAGGGATG	CAGGGGTGCCTGCACTAAG
KIAA1324_ex11	CAACTTGTCTGTGTGGCAGC	AAGAACTGAGGGTGTGCAGG	WNT2B_ex6	TAAGGGTATCCAGGGCAGC	GAAGGTGTGCATGCAAGG
KIAA1324_ex12	CCTCCCTTCCAGAGTGTCATC	GGATTCGGATATGGCCAAG			
KIAA1324_ex13	TGGGCTCTTCAATAAGCACTG	CAGATGGGGCTAACCCTCC			
KIAA1324_ex14	GTGATGGCACCTGGGATG	CCTCCTGGGTTCAGAGGG			
KIAA1324_ex15	AGAGGCCTCAATTGTTGCTG	TCTTCTGAGGTCATCCAAGC			
KIAA1324_ex16	AAAACTGCTACTCTTGGCTGC	CTGATCCATGGGCCTCC			
KIAA1324_ex17	GCCTTCCCTATTCCCTTCTG	CTCATGTGGGGCCTTTTACC			
KIAA1324_ex18	TGACAGGTATTGGAAGAATTGAAG	GCAGCAGAAATTATGGAGGAG			
KIAA1324_ex19	TGGTGGCTGTCCAATAACAC	GGATGGAATGACTTCAAGGG			
KIAA1324_ex20	GAGACAACAGGCTTCCAACC	GAACTCAGATCTCCCATTCCC			
KIAA1324_ex21	GGCACTTCCATGAGAGAAGG	CCCTGATCAATAATCCAACCC			
KIAA1324_ex22	AACCTCTCTACCTCTTTCCCC	AGCAGTGTCTCCTCCAGCC			
KIAA1324_ex23	CTTTGTCCCCAAACTTGCC	AGGCTTGCAAAGGTGCTATG			
LRIG2_ex1	GCCGATCCTCCTTTTCTAGC	CCCCTGTAGAGGGAAGGC	LRIG2_ex12	TGGCAAGAATAATGGATAGGTG	TCCTAGTGGAAAACCATTCTGG
LRIG2_ex2	TTGCCAGGTGAGTCTTGAAC	CAACCCAATGCCAAGACAG	LRIG2_ex13.1	GCCACTTTGTAAACAGAACACAC	TATTCCAGAGCTTCTCCAGC
LRIG2_ex3	GCCAGAAGGTAGAATCCAAGAG	CTGCTGCCACTGATCCAC	LRIG2_ex13.2	GCAGTGAGCAGCAGTGATTC	AAGTGGGCAGACTGACTAATCC
LRIG2_ex4	GCCAGAAGGTAGAATCCAAGAG	CCACAACAGTCCCCATAAAC	LRIG2_ex14	GGGGTGTCTTGCCTACTAGATTC	GACCAAGAGTAGCTCCCAAAG
LRIG2_ex5	TTTTGGCAGCTTTCTAGTGG	CAGGAAATTGGTTTGCTAAAATG	LRIG2_ex15.1	TTGGTCTAATTCAGTAGAAATTAGGG	TCATGCCCAATGCTACTCTG
LRIG2_ex6	GATCTTCAAGCTGCCTCACC	ACCCTAGCCATTCGACACTG	LRIG2_ex15.2	AGAACGACATTTCTTTGCTGC	TGCTACTATGACCAGCTCTCAG
LRIG2_ex7	AAGCTAGTTATTGAACTGCCTTG	TGATTTATTCCAATCTGCTCTAGTC	LRIG2_ex16	GCAAAATAGTCTTCTGAGTGTTGG	TTGCTTCAGATTAATGGAGAAACTAC
LRIG2_ex8	TTCCCATGGGCAGTAATGAG	TGCTCTGAGGCATAGGACAG	LRIG2_ex17	TGAAGAAAACTGACGATTGGC	TGCTTGCTTCCTGTCAAATG
LRIG2_ex9	TGCTCAATGTTGTGATGGTTG	TCATATTTACAGCCGGTTCTTG	LRIG2_ex18	TGTGTGTGATTTCAGTGTCCC	AAAGCAAATGCTGACTTCGG
LRIG2_ex10-11	AAAAGTACGCTACGCTTATTTTCAC	ACTTGGGTCACTCCCATCAC			

Abbreviations

AA-AMP	aminoacyl adenosine monophosphate
ADID	autosomal dominant intellectual disability
AMP	adenosine monophosphate
Amp^r	ampicillin resistance
APS	ammonium persulphate
ARID	autosomal recessive intellectual disability
ARS	aminoacyl-tRNA synthetase
ATP	adenosine-5'-triphosphate
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
Chr	chromosome
Cm^r	chloramphenicol resistance
CMTRIB	Charcot-Marie-Tooth disease, recessive intermediate , type B
CNS	central nervous system
CNV	copy number variation
CSF	cerebrospinal fluid
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
ddH_2O	double distilled water
ddNTP	dideoxynucleotide
DEPC	diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EYFP	enhanced yellow fluorescent protein
EtBr	ethidium bromide
FBS	fetal bovine serum
FW	forward

FXS	fragile X syndrome
GFP	green fluorescent protein
GST	glutathione S-transferase
GTP	guanosine-5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horseradish peroxidase
IPTG	isopropyl β-D-1-thiogalactopyranoside
ID	intellectual disability
IQ	intelligence quotient
kb	kilobase
kDa	kilo Dalton
Km^r	kanamycin resistance
LB	Lysogeny broth
LOD	logarithm of odds
LTP	long term potentiation
Mb	megabase
MIM	Mendelian Inheritance in Man
MR	mental retardation
MIR	magnetic resonance imaging
mRNA	messenger ribonucleic acid
NAD^+	nicotinamide adenine dinucleotide (oxidized form)
NADH	nicotinamide adenine dinucleotide (reduced form)
NCBI	National Center for Biotechnology Information
NGS	next generation sequencing
NLS	nuclear localization signal
NMD	nonsense-mediated mRNA decay
Nm^r	neomycin resistance
NSID	non-syndromic intellectual disability
OCF	occipitofrontal circumference
OD	optical density
OMIM	Online Mendelian inheritance in man
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBST	phosphate buffered saline with Tween 20
PCR	polymerase chain reaction
PDB	protein data bank
PFA	paraformaldehyde

Pfam	protein family
PP_i	pyrophosphate
P/S	penicillin-streptomycin
PVDF	poly-vinylidene difluoride
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT	room temperature
RT-PCR	reverse transcription PCR
RV	reverse
SSA	succinic semialdehyde
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SID	syndromic intellectual disability
SNP	single nucleotide polymorphism
TAE	tris-acetate-EDTA
TCA	tricarboxylic acid
TEMED	N,N,N',N'-tetramethyl-ethane-1,2-diamine
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
Tween 20	polyethylene glycol sorbitan monolaurate
TY	tryptone, yeast extract
UCSC	University of California Santa Cruz
UTR	untranslated region
v/v	volume/volume ratio
WES	whole exome sequencing
WGS	whole genome sequencing
w/v	weight/volume ratio
WHO	World Health Organisation
WT	wild-type
XLID	X-linked intellectual disability
X-Gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside

Three-letter and single-letter code of amino acids

Amino acid	Three letter	Single letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Selenocysteine	Sec	U
Pyrrolysine	Pyl	O